GEOGRAPHIC VARIATION IN MEDICARE SPENDING AND THE HEALTH, FUNCTIONING, AND BEHAVIORAL RISK FACTORS OF OLDER AMERICANS

by
Kurt Richard Herzer, MSc

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

January, 2016

ABSTRACT

Variation in healthcare delivery is ubiquitous in the United States. Two fundamental questions in health policy discussions today are these: what determinants explain the marked variation in healthcare spending across hospital referral regions in the United States and whether or not greater healthcare spending improves patient outcomes. This dissertation extends prior research by (1) examining the role that health behaviors and modifiable risk factors play in explaining differences in Medicare spending across regions and (2) evaluating how differences in spending affect beneficiaries' physical, cognitive, and mental health and functioning.

Dissertation aims were addressed using data from the Health and Retirement Study (HRS), a nationally representative longitudinal survey of older Americans, and this was linked to Medicare claims and regional spending characteristics from *The Dartmouth Atlas of Health Care*. Medicare beneficiaries age 65 or older and enrolled in traditional (fee-for-service) Medicare comprised the study population. Price-adjusted Medicare spending was measured as spending for care in a variety of settings (inpatient, skilled nursing facilities, outpatient, physician office, home health, and hospice) as well as spending for durable medical equipment. Analytic methods included regression-based decomposition techniques and instrumental variables analysis.

In assessing the determinants of regional variation in spending, Medicare beneficiaries' observable characteristics (e.g., sociodemographics, behavioral risk factors, health status factors) collectively explained 17% of regional differences in price-adjusted Medicare spending. Behavioral risk factors, specifically—smoking status, alcohol consumption, body mass index, and physical activity—explained 7% of the difference in spending between higher- and lower-spending regions.

In examining the effects of Medicare spending on health and functional outcomes among hospitalized Medicare beneficiaries (after adjusting for confounding due to health status), a 10% increase in price-adjusted, 1-year post-admission Medicare spending was associated with reductions in the probability of new limitations in instrumental activities of daily living (-1.96 percentage points; 95% confidence interval [CI], -3.88 to -0.03; P=0.05), new depressive symptoms (-2.31 percentage points; 95% CI, -4.04 to -0.59; P=0.009), and 1-year mortality (-2.02 percentage points; 95% CI, -3.57 to -0.46; P=0.01). There was no association between higher Medicare spending and self-rated health status, limitations in activities of daily living, pain, or cognitive functioning.

This dissertation provides policymakers with new information about the importance of behavioral risk factors as determinants of regional variation in Medicare spending and the impact of healthcare spending on multiple dimensions of health and functioning in Medicare beneficiaries.

Advisor: Albert Wu, M.D., M.P.H.

Readers: Peter Pronovost, M.D., Ph.D.
Lauren Nicholas, Ph.D., M.P.P.
Roland Thorpe, Ph.D.

Alternates: Judith Kasper, Ph.D.
Marlis Gonzalez-Fernandez, M.D., Ph.D.
Evan Mayo-Wilson, D.Phil.

ACKNOWLEDGEMENTS

Completing this dissertation would have been impossible were it not for many individuals whom I wish to thank briefly.

First and foremost, I am grateful to my thesis advisory committee: Albert Wu, Lauren Nicholas, and Peter Pronovost. I first met Albert in Geneva, Switzerland, in 2008 while he was on sabbatical at the World Health Organization (WHO) and I was working there for an extended summer. Albert was my supervisor at WHO and little did I know then that years later he would be my Ph.D. advisor. As the dissertation took shape, Albert entertained many of my meandering ideas and always maintained an enthusiasm and positive focus that was invaluable to my progress. Albert is a wonderful mentor and teacher who has helped me keep my dual training in medicine and health policy in perspective.

I am especially grateful to Lauren, who repeatedly challenged me to explore my data deeper and who set an example of excellence in research. This dissertation would not have been possible without her generosity in sharing both the data and her time. Because of Lauren, I have had the opportunity to learn and work with methods I never would have otherwise; the time she invested in mentoring me has made me stronger methodologically, which I believe will serve me well over my career.

I would undoubtedly not have undertaken a dual degree training had it not been for Peter, who encouraged me to apply to M.D.-Ph.D. programs when I was a junior in college and who has relentlessly supported me for more than a decade. Peter taught me about leadership and vision, and the pursuit of quality and excellence in all domains of life.

I also wish to acknowledge Roland Thorpe, Judy Kasper, Karen Davis, Jodi Segal, Lynette Mark, Marlis Gonzalez-Fernandez, and Evan Mayo-Wilson who served on, or graciously offered to serve on, my various oral examination committees.

I am grateful to all of my other mentors, particularly Lynette Mark. After a serendipitous encounter when I was still a high school student, Lynette helped develop my early academic interests in healthcare quality and safety and provided an example of clinical excellence. She has been an unwavering source of support to me as I matured as a researcher and chose to go to medical school. I am also deeply grateful to Susan Turley, a dear friend, who has taught me much about academic writing and provided feedback on countless papers and documents over the years.

I am indebted to the Medical Scientist Training Program/M.D.-Ph.D. Program for both its financial assistance and its unqualified support of my path to pursue a Ph.D. in health policy rather than the basic sciences. I am particularly grateful to Robert Siliciano, Andrea Cox, and Sharon Welling for their leadership of the program, as well as Bernadine Harper and Martha Buntin, who worked behind the scenes to make my dual degree training possible. I am also thankful for the friendships engendered by this program, especially those with Vivek Charu and Melissa Liu.

In the Department of Health Policy and Management, I wish to thank Mary Sewell for helping to coordinate my exams, Keri Bowers for her assistance in managing my grant, and the leadership of the department for its support of Ph.D. students. I am also grateful to the leadership and staff of the Hopkins Population Center, who provided the logistical means to access and work with the restricted data used in this dissertation—specifically, Michele Trieb and Jennifer Darragh, from whom I have learned an immense amount about working with restricted data.

This research was funded by the National Institute on Aging, under Award Number R36AG051727, as well as the Health Assessment Laboratory/Medical Outcomes Trust. I am grateful to both organizations for their support of my dissertation research. This dissertation does not represent the official views of either organization.

Several additional individuals must be acknowledged: my grandparents, whose sacrifices and hard work made my life choices possible; my parents, Karl and Patrice Herzer, who never placed an upper limit on my aspirations; and my cousin, Alexandra Herzer, in whom I have always found a kindred spirit and steadfast companion. Lastly, I thank my wife, Amber Herzer, for keeping me grounded and centered on what matters most in life, and for her support during my protracted dual degree training.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

BACKGROUND

The current and perennial debate about variation in regional healthcare spending in the United States had its origin in an unexpected discovery many years ago. In 1938, J. Alison Glover discovered that the rate of tonsillectomies among British school children varied widely across localities in England and Wales.[1] It was not until 40 years later that this same phenomenon, coupled with the recognition that differences in physicians' medical practices were largely responsible for the varying tonsillectomy rates, was noted by John Wennberg and Alan Gittelsohn to exist in Vermont.[2] Since Wennberg's pioneering work describing small area variation in the 1970s,[2] an abundance of research has been conducted relating to geographic variation in healthcare delivery. In 2013, publication of the Institute of Medicine's (IOM) landmark report, *Variation in Health Care Spending: Target Decision Making, Not Geography,* reinvigorated public interest in the relationship between geography and healthcare delivery and spending.[3] Two fundamental questions that are a part of health policy discussions today are these: what causes the marked variation in healthcare spending across the 306 hospital referral regions (HRRs) in the United States and whether or not greater healthcare spending improves patient outcomes.[4-8] This dissertation addresses both of these questions.

With passage of the Affordable Care Act in 2010, an influx of aging Americans into Medicare, and recent projections that Medicare spending will continue to grow,[9] policymakers face renewed pressure to understand and curtail Medicare spending.[3,10] Variation in utilization has been viewed as an opportunity to reduce plausibly unnecessary spending; in fact, some experts estimate that 30% of healthcare spending could be eliminated if higher-spending regions reduced the amount of care they deliver to the level of lower-spending regions of equal quality.[10-12] Nevertheless, substantial debate still surrounds the causes and consequences of variation as well as the requisite policy interventions.[13] That debate has been framed by a handful of published studies that principally assessed mortality as an outcome and used methods and administrative datasets that restricted causal inference. Despite their equivocal results about whether more spending improves health, these articles have had—and continue to have—a significant impact on healthcare policymaking.

This dissertation asserts that a deeper understanding is needed of the factors contributing to regional variation in healthcare spending and the relationship between healthcare spending and outcomes. First, this deeper understanding is needed because the impact of health behaviors and modifiable risk factors on variation in spending and utilization is largely unknown. Smoking, alcohol consumption, obesity, and limited physical activity—all of which are prevalent within the rapidly aging population—are prominent examples of these patient determinants of utilization.[14] Despite recognition that behavioral risk factors vary geographically and have contributed substantially to the growth of healthcare spending over the past 20 years,[15-17] these factors have not been adequately assessed as possible drivers of regional variation in spending.[18] Second, the current literature is constrained by its reliance on mortality as the primary measure of health; yet many costly and discretionary treatments mainly impact quality of life, rather than length of life.

This dissertation seeks to advance this current body of evidence by (1) studying a broad set of behavioral risk factors in addition to health outcomes, (2) using rigorous econometric methods to estimate causal effects and address unobserved confounding, and (3) leveraging the unique opportunities afforded by accessing data in the Health and Retirement Study (HRS), a nationally-representative longitudinal survey of older Americans that is linked to Medicare claims.

DISSERTATION AIMS

The aims of this dissertation were to:

1. Examine the literature on the determinants of healthcare spending and geographic variation in spending as well as evidence of the relationship between the intensity of healthcare spending and patients' outcomes;
2. Develop a deeper understanding of the role of behavioral risk factors in explaining differences in spending and utilization across regions; and
3. Examine the effects of healthcare spending on the physical, cognitive, and mental health and functioning outcomes of hospitalized Medicare beneficiaries.

ORGANIZATION OF THE DISSERTATION

This dissertation consists of five chapters. This chapter*, Chapter 1*, introduces the dissertation topic, enumerates the aims of the dissertation, and describes the structure and format of the dissertation, as contained in subsequent chapters.

Chapter 2 reviews the literature and, from this, develops a conceptual framework that informs subsequent analyses (Aim 1).

Chapter 3 and *Chapter 4* comprise the main empirical manuscripts of the dissertation and address the determinants and consequences of geographic variation in healthcare spending, respectively. For both manuscripts, data from the HRS were linked to Medicare claims and regional spending characteristics from *The Dartmouth Atlas of Health Care*.[19] More specifically, *Chapter 3* investigates the role of behavioral risk factors as possible explanatory factors of geographic variation in Medicare spending (Aim 2). The difference in cumulative spending on Medicare beneficiaries between 2004 and 2006 in higher- versus lower-spending regions was examined using a regression-based decomposition technique originally developed in the labor economics field. This method assesses whether regional differences in Medicare spending derive from observable differences in beneficiary characteristics, such as their behavioral risk factors, or whether these differences derive from provider practice patterns that treat similar beneficiaries differently. *Chapter 4* examines the association between Medicare spending during and after acute hospitalization and beneficiaries' subsequent health and functioning (Aim 3). Instrumental variables analysis, a common econometric technique, was used to reduce confounding related to beneficiaries' health status and severity of illness and to estimate causal effects of spending differences on self-rated health, functional status, pain, cognition, depressive symptoms, and mortality. Both manuscripts contain appendices detailing extensive additional analyses that were designed to evaluate the robustness of the main results. The main findings of both manuscripts were supported by these sensitivity analyses.

Chapter 5 summarizes the key findings of the dissertation, reiterates their policy implications, offers recommendations for future research, and concludes the dissertation. When viewed as a whole, these manuscripts provide policymakers with important, new information about the determinants and consequences of geographic variation in Medicare spending.

CHAPTER TWO: LITERATURE REVIEW (MANUSCRIPT #1)

DETERMINANTS AND CONSEQUENCES OF GEOGRAPHIC VARIATION IN HEALTHCARE SPENDING: A REVIEW OF THE LITERATURE AND A PROPOSED CONCEPTUAL FRAMEWORK

by

Kurt Richard Herzer

ABSTRACT

Two fundamental questions in health policy are what causes differences in spending across regions and whether higher healthcare spending improves patient outcomes. This manuscript provides a narrative review of the literature pertaining to the determinants and consequences of geographic variation in healthcare spending. Demand-side and supply-side determinants of spending and of regional variation in spending are reviewed, with an emphasis on those factors particularly salient to the subsequent manuscripts that comprise this dissertation. Studies assessing the effects of healthcare spending on outcomes are heterogeneous, and the evidence that additional spending improves outcomes is equivocal. Confounding by unobserved severity of illness and the predominant use of mortality as the primary representation of health status are important limitations of this literature. Based on this literature review, Ronald Andersen's widely known behavioral model of health services utilization is then selected, adapted, and applied to the study of geographic variation in healthcare spending. Specifically, this proposed conceptual model diagrams the pathways through which elements of the healthcare system and predisposing characteristics of the population influence healthcare utilization and outcomes. Future directions and challenges for research conclude the manuscript.

INTRODUCTION

Geographic variation in healthcare use across the United States is pronounced, pervasive, and persistent.[3] Medicare beneficiaries living in high-spending regions receive significantly more healthcare than those in low-spending regions.[2,10] The Institute of Medicine's (IOM) landmark 2013 report, *Variation in Health Care Spending: Target Decision Making, Not Geography*, returned the topic of practice variation to the fore of health policy. Two fundamental questions in health policy are what causes the marked variation in spending across the 306 hospital referral regions (HRRs) in the United States and whether greater healthcare spending improves patient outcomes.

The objective of this manuscript was to critically review and synthesize pertinent literature pertaining to this dissertation and to inform the empiric manuscripts that follow. The organization of this manuscript corresponds to the 3 Specific Aims described previously: determinants of geographic variation, with a focus on distinguishing demand-side versus supply-side factors (Part 1), and consequences of geographic variation (Part 2). Following Part 2, a conceptual framework is presented that integrates this literature.

METHODS

To identify relevant literature for this review, searches were conducted using PubMed, Google Scholar, and, to a lesser extent, EconLit. Given the likelihood of wide-ranging gray literature related to this topic, searches were also conducted through specific organizations: Kaiser Family Foundation, Centers for Medicare and Medicaid Services (CMS), Congressional Research Service, American Hospital Association, Robert Wood Johnson Foundation, and the IOM. Key search terms included the following: *geographic variation*, *regional variation*, *practice variation*, *healthcare spending*, *healthcare utilization*, *treatment intensity*, *Medicare*, and *outcomes*. Numerous reviews of the geographic variation literature exist,[13,20-23] including the recent IOM report that was extensive in its coverage.[3] These existing reviews provided a basis for structuring this manuscript. The bibliographies of these reviews were also examined and references were extracted with the belief that these references would be emblematic of key literature.

PART 1. DETERMINANTS OF GEOGRAPHIC VARIATION IN SPENDING AND UTILIZATION

The causes and determinants of geographic variation in healthcare spending have been a source of growing interest in the literature. A recent exhaustive review of the subject by Manning and colleagues (2012)[18] (as a supplement to the IOM's report on geographic variation) informs the focused review that follows. The determinants of variation in utilization can be broadly dichotomized into demand-side and supply-side factors.[24,25] We discuss both of these sets of determinants of utilization and their propensity to vary regionally. Pertinent issues related to defining or measuring these determinants are also discussed.

Demand-side Factors

Demand-side factors relate to how and why patients seek health care, and they include underlying health status, sociodemographic characteristics (including access to health insurance coverage), health behaviors and modifiable risk factors, and patient preferences related to health care. These factors are typically accounted for as independent variables in regression models.

Health Status

Health status of the patient is a crucial determinant of demand for health care.[26] Sicker patients require (and receive) more care and incur greater spending and utilization. As such, it is vital to adjust for health status when comparing the healthcare utilization of different patients, and when comparing utilization of patients in different regions because health also varies geographically.[27,28] For example, using data from the National Health Interview Survey, Kachan and colleagues (2014)[27] found significant state-level regional differences in health-related quality of life among older Americans. Other studies have demonstrated that functional status[29] and life expectancy[30] also vary geographically. The empirical challenges posed by health status as a confounder are discussed in Part 2 of this literature review. Here, discussion centers on how health status is measured and how it contributes to regional variation in healthcare spending and utilization.

Health status can be measured in a variety of ways. For example, patients' medical diseases are commonly elicited from administrative records, survey data, or electronic health records. Comorbid diseases can be identified using the *International Classification of Diseases*, Ninth Revision (ICD-9) or 10th Revision (ICD-10) codes present in administrative records, and these codes are also used in classification systems such as Medicare's Hierarchical Condition Categories,[31] the Charlson Index,[32] or Elixhauser comorbidities.[33] Comorbid diseases commonly controlled for in analyses include the following: hypertension, diabetes, coronary heart disease, stroke, cancer, and acute myocardial infarction.[18] Patients can also self-report health status, and such self-reporting may be less subject to concerns about bias related to physician diagnostic practices.[34] Finally, it is not uncommon for health status variables to be measured at the population level (such as incidence of disease or mortality rates within specific regions or geographical areas) when data for these variables are not available for individuals.

Despite consensus that health status is a key driver of healthcare utilization, debate persists amongst scholars about the importance of health status for explaining regional variation in utilization. That debate is framed by two divergent viewpoints: the first is the belief that regional differences in utilization are really due to underlying (unmeasured) differences in patient health status and that better risk adjustment would attenuate unexplained differences; the second is the belief that regional differences in utilization are real, cannot be explained away solely by case mix, and that supply-side factors such as physician aggressiveness are the key determinants.[35] Consequently, studies have come to divergent conclusions about the extent to which health status accounts for variation in utilization across regions.[36,37] This variability is a result of these studies using different methods to control for health status and using different geographic units of analysis.

Two studies, relying on the same data source (the Medicare Current Beneficiary Survey [MCBS]) provide insight into the extent to which health status can explain regional variation in healthcare spending. Sutherland and colleagues (2009)[38] used MCBS data (at the patient level) and included blood pressure, diabetes, body mass index, smoking history, and self-rated health as health status measures; they found that these variables explained 18% of the regional differences in spending between the highest and lowest quintiles. Subsequently, Zuckerman and colleagues (2010)[37] used the MCBS but added over 10 additional health variables; they found that health status factors explained 29% of the differences in spending between regions in high versus low Medicare spending quintiles. The recent publication of a paper by Reschovsky and colleagues (2013)[39] that implemented a modified version of Medicare's Hierarchical Condition Categories model (by including only diagnoses that were less susceptible to physician discretionary behavior) concluded that population health could explain more than 75% to 85% of spending variation. However, the Reschovsky study did not directly address the concern that there is a lower threshold for diagnosis in higher intensity regions and that use of the Hierarchical Condition Categories model may over-adjust regional differences in spending.[40,41]

The intensity of patient observation varies regionally, and diagnoses recorded in medical records and claims do not exclusively reflect underlying disease burden; as such, the use of clinical or claims-based diagnoses for risk adjustment may introduce additional biases.[42,43] Song and colleagues (2010)[42] exploited a natural experiment in which patients were moving to higher-versus-lower intensity regions; they demonstrated that those patients moving to higher-intensity regions experienced more physician visits, diagnostic tests, and imaging exams—and accrued more diagnoses—following their move, but with no differences in their 3-year post-move mortality. Similarly, Welch and colleagues (2011)[43] demonstrated a positive relationship between regional intensity of patient observation (measured as frequency of physician visits, diagnostic tests, laboratory exams) and diagnoses of chronic illnesses, as well as an inverse relationship between the regional frequency of diagnoses and case fatality from chronic disease.

In light of this evidence that diagnostic practices vary regionally, it is evident that patients in higher-intensity regions accumulate more diagnoses and thus appear sicker. However, this is an artifact of diagnostic differences rather than differences in underlying health status, suggesting a judicious application of risk adjustment approaches in studies related to geographic variation. Recent work by John Wennberg and colleagues (2013)[44] compared a standard method of risk adjustment to a visit-corrected method that specifically adjusts for bias induced by observational intensity using HRR-level data on the frequency of physician visits as a proxy for intensity of observation. They compared the standard method to the visit-corrected method for three common risk adjustment schemes: the Charlson Index, the Iezzoni chronic condition count, and the Hierarchical Condition Categories risk score. They found that the visit-corrected method reduced the bias associated with observational intensity. David Wennberg and colleagues (2014)[41] developed two new approaches to health risk adjustment—a poverty index and a population health index (consisting of self-reported illness, obesity, smoking status, and the regional incidence of admission to the hospital for hip fractures and strokes). Both approaches outperformed the standard Hierarchical Condition Categories index in terms of explaining and reducing variation in age-sex-race adjusted mortality and in terms of not exhibiting bias due to observational intensity. A final limitation of using diagnoses recorded in administrative data as risk adjusters is that they do not discriminate between different disease severity levels, even among patients with the same diagnosis, and so concerns related to omitted variables bias remain (this source of omitted variables bias is discussed further in Part 2 of the literature review).[45]

Sociodemographic Characteristics

Sociodemographic characteristics—age, sex, race and ethnicity, education, and income or wealth—are associated with health status and healthcare utilization. It is well recognized that advancing age is associated with increased utilization of hospital and physician services and pharmaceuticals, secondary to chronic diseases and functional declines.[46] Nevertheless, greater use of effective therapies, such as antihypertensive drugs, may reduce downstream utilization for complications of untreated or undertreated chronic diseases. As such, the independent effect of age on utilization is not clear-cut.[46] Sex is associated with different care-seeking behaviors and perceptions of illness,[47] social roles, and prevalence of illnesses. Some studies have identified that women utilize more healthcare services than men;[48] but conversely, other studies have found similar utilization in men and women after women's childbearing years and men's higher age-specific mortality rates were controlled for.[49] There are marked disparities in the U.S. in both the quantity and quality of health care received by racial and ethnic minorities.[50] Regional patterns of racial disparities in healthcare utilization rates have also been documented.[51-53] Age, sex, and race are frequently controlled for when studying health expenditures or outcomes of interest. For example, the Dartmouth Atlas of Health Care adjusts its measures for age, sex, and race to capture differences in beneficiary characteristics across regions.[54] These same demographic characteristics often also serve as proxy measures for health status when health status variables are not available.

Factors such as education, income, and wealth and can affect decisions to access and utilize health care (and vary demographically and geographically).[18,55] Socioeconomic status, whether measured by education or income, is associated with healthcare utilization and health behaviors as well as morbidity and mortality.[56] Education can improve health through better occupational factors and work conditions, and an increased knowledge about lifestyle factors, health behaviors, and social-psychological resources.[57] Greater income can improve access to health insurance and consequently to health care;[58] it can also provide for better living environments, educational opportunities, and recreation, all of which tend to enhance health status. Because measuring income accurately poses a challenge (owing to reporting biases) and only provides a cross-sectional view of an individual's financial resources, wealth can be used an alternative measure of socioeconomic status that is also significantly associated with health,[59] but not necessarily with utilization.[60] Income or wealth also represent important variables to control for when studying healthcare utilization and may also act indirectly as proxies for other determinants of health.[61] Insurance coverage status is related to income to the extent that access to different types of health insurance or health insurance programs is often related employment or income levels (e.g., access to Medicaid as a form of coverage for lower income Americans as opposed to private insurance coverage for employed or higher income Americans). Health insurance in turn affects access to care,[62] and access is known to influence spending and utilization. Fisher and colleagues (2003)[5] found that residing in a higher-spending region was associated with worse access.

In sum, differences in how sociodemographic variables are measured and coded and which are included or excluded—all of these preclude clear conclusions about the significance of these factors in explaining regional variation in spending.[18] The IOM report found that controlling for age and sex had a minor effect on geographic variation in spending; similarly, race and income had minor effects once health status was accounted for in analyses.[3]

Health Behaviors and Modifiable Risk Factors

Health behaviors and modifiable risk factors, such as smoking, drinking, obesity, and exercise affect utilization of health care and are directly correlated with health.[63] Smoking is associated with increased visits to specialists and hospitalizations.[64] Heavy alcohol users disproportionately utilize emergency and acute healthcare services[64,65] Compared with nonobese patients, obese patients incur more visits to primary care physicians and specialists, use more diagnostic services, and have greater charges in aggregate and across a range of services.[66]

Individuals who smoke, drink alcohol in excess, have unhealthy diets, are obese, and are physically inactive are at greater risk of developing chronic diseases, such as heart disease, stroke, cancer, and diabetes.[67-69] Smoking increases the risk of numerous diseases, including cardiovascular and cerebrovascular disease (such as, acute myocardial infarction and stroke), respiratory diseases (such as, chronic obstructive pulmonary disease and asthma), cancers (particularly lung cancer[70]), and diabetes, among many others.[71,72] Alcohol consumption at light-to-moderate levels may be associated with health benefits, such as decreased risk of cardiovascular diseases (though what constitutes “moderate” varies across studies).[73-75] However, heavy drinking is associated with increased risk of cardiovascular disease, stroke, cancer, gastrointestinal diseases, and the health consequences of injuries. [76-80] Field and colleagues (2001)[81] followed individuals for 10 years and found that obesity was associated with increased risk of diabetes, gallstones, hypertension, heart disease, colon cancer, and stroke.

Once acquired, the chronic diseases caused by or linked to these behavioral risk factors can precipitate healthcare utilization related to treatment and management.[82] For example, in 2005 the U.S. spent $190 billion on healthcare expenses related to obesity.[83] These same chronic diseases are also significant sources of morbidity and mortality in the U.S., especially among older Americans.[14,84] Mokdad and colleagues (2004),[84] in revisiting a seminal 1993 article by McGinnis and Foege,[85] found that tobacco, alcohol consumption, and poor diet and physical inactivity accounted for nearly half of all deaths in the U.S. in 2000.

There are a variety of ways to define and measure behavioral risk factors. For example, smoking is often measured as duration in years (current age minus age since first smoked regularly), elapsed time since quitting for past smokers (age since last smoked minus age since first smoked), and pack-years (number of years smoked multiplied by the number of packs per day). Common approaches for measuring alcohol consumption include the quantity-frequency approach[86] (number of drinks per day times the number of drinking days per week) and the CAGE screening assessment (2 or more positive responses to a 4-item instrument suggests alcohol abuse/dependence).[87] Physical activity is often queried by asking individuals their frequency, duration, and intensity of exercise, often within a specific interval of time (such as a week).[88] Obesity is typically measured using body mass index (BMI) classification; for example, a BMI between 25 and 29.9 is considered overweight and greater than 30 is considered obese. As described by Riekert and colleagues (2014),[88] the measurement of health behaviors can be objective or subjective. Objective and direct measurement (as the behavior happens) occurs with behavioral observation or electronic monitoring. Objective and indirect measurement (after the behavior has occurred) occurs with biochemical analysis. Subjective and indirect measurement occurs via different forms of individual or proxy ratings of past behavior. Self-reporting and rating of health behaviors is common in large and longitudinal surveys because it facilitates data collection, but, because it does not directly measure behaviors, it is subject to reporter and recall bias.

Behavioral risk factors also vary by demographic characteristics, socioeconomic status, and geography.[15] For example, Victor Fuchs acknowledged years ago that Utah has a relatively low prevalence of smoking and drinking, characteristics he attributed to the high proportion of Mormons residing in the state who disavow these practices.[28] Few studies in the geographic variation literature have included measures of behavioral risk factors and the importance of these factors in explaining variation in healthcare spending and utilization remains uncertain.[18]

Patient Preferences

Patient preferences related to health care can also influence healthcare utilization. If patient preferences vary regionally, then including measures of preferences would be analytically important.[18] For example, if patients in certain regions of the country demand more office visits, greater use of intensive care, more diagnostic tests and imaging, and prefer to see specialists, then supply of these resources, such as physicians, beds, and imaging equipment, would increase to satisfy the preference for greater intensity of care.[89]

Preferences are defined differently across studies, and few datasets collect measures of patient preferences across an array of geographic areas. Dartmouth researchers have defined preference-sensitive care as, “interventions for which there is more than one option and where the outcomes will differ according to the option used.”[90] Fisher and Wennberg (2003)[91] argue that preference-sensitive care is a problem of misuse (rather than underuse or overuse) and occurs when medical theory is strong but actual medical evidence (such as from experimental or observational studies) is variable or ambiguous. As such, preference-sensitive care involves two or more treatment alternatives and necessitates tradeoffs between different sets of risks and benefits. Elective surgery is archetypical of preference-sensitive care. Ideally, treatment selection would be based on informed patient decision-making buttressed by the best available information; in practice, physicians’ opinions about outcomes and their beliefs about patient preferences likely drives the decision.

While the extent to which patient preferences explain variation in spending or utilization remains unclear, limited empiric research suggests that patient preferences account for relatively little of the variation in spending across regions. Baker and colleagues (2014)[92] found that patients' preferences explain 5% of the variation in Medicare spending across hospital referral regions. Using national survey data, Anthony and colleagues (2009)[89] found significant variation among Medicare beneficiaries in their preferences for utilizing health care, but the distribution of preferences was similar across regions. Examining EOL care, Barnato and colleagues (2007)[93] failed to identify a significant relationship between regional expenditures for EOL care and patient preferences. Recent work by Cutler and colleagues (2013)[94] demonstrated that physician preferences were significantly more important for explaining geographic variation in utilization when compared with patient preferences.

Supply-side Factors

The geographic variation literature has emphasized supply-side factors (those attributable to physicians and provider organizations) as the important determinants of geographic variation. As such, many proposed policy interventions to curb unwarranted variation are inculcated with a belief that the greatest opportunities to reduce variation in utilization and improve quality can be achieved by influencing how providers deliver care. This focus underlies such initiatives as pay-for-performance, value-based purchasing, and accountable care organizations. Supply-side determinants exist at multiple levels and include clinician factors, hospital factors, and regional factors; these levels are not necessarily mutually exclusive because of their nested structure (i.e., clinicians work within hospitals, and hospitals define important aspects of the region wherein they operate). Healthcare prices are also an important supply-side determinant of geographic variation in healthcare spending.

Clinician Factors

Clinician factors refer to determinants of utilization at the individual provider level. Clinician—namely physician—factors are generally agreed to account for a significant fraction of healthcare costs. As an occupational group, physicians constitute a fraction of 1% of the U.S. labor force, but collectively they influence roughly 17% of national GDP by directing how over 90% of every healthcare dollar is spent.[95] Supplier-induced demand is offered as one explanation of how physicians influence healthcare expenditures. In a principle-agent framework, the physician (or other healthcare provider) who is incentivized to provide more services, can shift a patient's demand curve beyond what the patient would otherwise demand. However, the existence of large variations in utilization in countries that have national healthcare systems (and few incentives for physicians to use more care) potentially contradicts suppler-induced demand as a putative explanation of variation in utilization.

At the bedside, physicians may recommend different treatments due to differences in diagnostic skill, beliefs in treatments' efficacies,[94,96] tolerance of medical uncertainty,[97] medical school and residency programs attended, and cultural/social norms related to specific diseases and treatments.[98] For example, physicians' varying tolerances of clinical uncertainty may result in different predilections for ordering tests or for using procedures that directly affect utilization.[95,99]

In the absence of clear evidence about the best therapeutic option, local norms and institutional culture may emphasize one alternative over others.[100] Burke and colleagues (2010)[101] and de Jong and colleagues (2003 and 2006)[102,103] identified that professional interactions amongst providers are important factors in guiding care decisions. For example, physicians who train in the same programs are more likely to share similar treatment patterns compared with physicians who train in different programs.[94,96,104] These practice patterns change slowly and may not evolve over time to a common norm.[105] In addition, providers (perhaps, especially trainees) may emulate the practices of peers, superiors, or opinion leaders when deciding among treatment approaches.[106] Clinical practice guidelines and decision aids have been proposed as interventions to standardize knowledge and practice and to reduce practice variation, but there is little evidence that these approaches alone can achieve these aims.[107,108]

The main question of import, however, is whether these factors vary across regions and whether physician preferences agglomerate geographically. While empiric work addressing this question remains limited, the recent study by Cutler and colleagues (2013)[94] (mentioned previously), used strategic surveys of physicians and patients and demonstrated that physicians' preferences were geographically correlated and outweigh patients' preferences in explaining regional variation in utilization. Underlying physicians' treatment preferences were the physicians’ own apparently idiosyncratic beliefs about treatments' efficacies (which were often inconsistent with professional practice guidelines); this was in addition to physician-perceived pressure to satisfy patients or referring physicians.

Hospital Factors

Two of the most salient hospital factors that contribute to differences in spending are a hospital’s ownership status and its resource supply.[18] Ownership status of a hospital (i.e., whether a hospital is for-profit, not-for-profit, or government-operated) has been proposed as a source of variation in utilization. As suggested by Kenneth Arrow’s formative work (1963)[109] on information asymmetries, if patients are unable to discriminate differences in quality between hospitals, then a hospital’s for-profit status could be perceived as an imperfect proxy signaling lower quality. In Arrow’s model, lower quality stems from a profit-focused orientation that would undermine spending on quality improvement.[18] Newhouse (1970)[110] later proposed that for-profit and not-for-profit hospitals have different objectives: for-profit hospitals seek to maximize profits, whereas not-for-profit hospitals seek to maximize quality of care *and* patient volumes. This theory implies that regions with more for-profit hospitals could have greater inpatient spending and utilization in clinical areas that are more profitable. Some empirical analyses, such as an assessment by Sloan and colleagues (2001),[111] found that for-profit hospitals were indeed more expensive than their not-for-profit or government-operated counterparts, but with no appreciable difference in quality.

Bed supply has also been viewed as a driver of variation in spending and utilization, especially for acute care beds.[112,113] Researchers at Dartmouth often describe such care as "supply-sensitive" because the supply of a resource (e.g., hospital bed) influences utilization. Supply-sensitive care is characterized by weak medical theory and medical evidence, a high per capita supply of the resource in question, and a variable importance of patient preferences.[91] Roemer's Law—a bed built is a bed filled—is often invoked by studies finding that hospitals with more beds have greater utilization. Roemer's law is at odds with the more conventional economic interpretation that believes inversely that supply follows demand. Manning and colleagues (2012) identify that there is a dearth of research on the causal chain underlying this association,[18] but several empirical analyses have produced results consistent with Roemer's Law.[114-116]

Regional Factors

Regional factors refer to factors best conceptualized at an area level because they either affect many providers and hospitals or are characteristics amalgamated over many providers or hospitals (e.g., practice norms, regulatory and legal climates, and financial incentives).[117] Area-level practice norms are epitomized by medical and surgical "signatures" for specific procedures, where physicians in specific areas use certain procedures at specific, stable rates that are unique to that area.[90,118] There are also differences across regions in the distribution or density of providers and their type. Studies have shown that regions characterized by greater numbers of medical specialists have associated higher spending, and, conversely, that regions with greater numbers of primary care providers have lower spending.[5,18,119] The existence of medical schools, academic medical centers, and research institutions within a region is possibly relevant because of the way in which such organizations set local performance standards and socialize and train providers who subsequently populate the local market.[120,121] David Molitor (2011)[98] demonstrated that cardiologists who migrate from one geographic region to other dissimilar regions (HRRs), practice more similarly to each other before the move versus after the move, suggesting that aspects of the first regional practice environment most strongly influence physician behavior. With this identification strategy, Molitor estimates that environmental factors are roughly twice as important for influencing physician behavior as physician-specific factors. In contrast, Grytten and colleagues (2003)[122] examined primary care physicians in Norway and found that their practice styles were stable before and after moving to different municipalities. The divergent conclusions suggested by these studies may relate to underlying differences in the study populations (specialists versus generalists) or the inherent differences between the United States and Norway. Nevertheless, the factors influencing physician utilization patterns are diverse, multifactorial, and incompletely understood.

Other characteristics operating at a regional level include malpractice environment, local emergency response systems, and decisions of local legislatures (e.g., certificate-of-need laws that affect purchase of expensive technology, capital expenditures, or medical school openings and closings).[18] As these topics are not directly germane to the dissertation, they are not discussed further.

Healthcare Prices

Lastly, prices vary geographically. For example, providers in different regions face different input costs, such as land prices and wage rates.[18] In prospective payment systems, such as inpatient care provided under Medicare, the price per episode for a specific diagnosis-related group

(DRG) reflects differences in input costs and market competition. Gottlieb and colleagues (2010)[54] demonstrated that utilization rather than local price differences drive regional variation in spending within Medicare, a conclusion buttressed by the IOM report. In contrast, in the commercial insurance market, regional differences in prices explain an estimated 70% of the total geographic variation in spending.[123] Price adjustment or standardization removes the effects of regional variation in input costs, such as those related to capital, labor, and overhead (rent and liability costs), and better reflects utilization of services.[124]

PART 2. CONSEQUENCES OF GEOGRAPHIC VARIATION: EVALUATING THE SPENDING-OUTCOME RELATIONSHIP

Hussey and colleagues (2013)[23] systematically reviewed the literature on the association between healthcare spending and patient outcomes, providing an important source of key references and topical organization for Part 2 of this literature review. In general, the literature on consequences of geographic variation in healthcare spending/utilization is heterogeneous; studies have used different data sources, levels of analysis (areas, providers, patients), dependent variables (quality measures of structure, process, or outcome), measures of costs or spending, and statistical methods (particularly with respect to adjusting for health status).[23] This part of the manuscript reviews the types of spending and outcome measures commonly used, methods to address confounding (specifically, how different study designs account for health status), and key findings from the existing literature.

Spending Measures

In discussing spending measures, it is important to note that studies of the spending-outcome relationship have used different levels of analysis with which to assess the effects of spending. Hospital-level studies are among the most common, followed by area- or region-level studies.[23] Area-level analyses are highly variable. Some are conducted exclusively at the area- or region-level (e.g., by comparing cross-sectional average spending in a state to that of average quality), as in the studies by Baicker and Chandra (2004)[119] and Cooper (2009).[6] Other studies, such as the classic articles by Fisher and colleagues (2003),[4,5] measured spending at the area-level (HRR) but examined patient-level outcomes. HRRs are commonly used regional/geographic designations in this literature. There are 306 HRRs in the U.S., which represent regional healthcare markets for tertiary medical care. They were originally defined as being the catchment areas around tertiary medical centers where patients were referred for major cardiovascular surgical and neurosurgical procedures.[19]

Studies have used different terminology to describe healthcare spending. For example, *expenditures*, *utilization*, and *costs* are, at times, used interchangeably to describe a similar underlying construct. *Expenditure*, as used in the literature, is typically measured as the rate paid for a service (price) multiplied by the number of services provided (quantity).[117] As such, regional variation in expenditures derives from regional variation in both price and quantity. Price-adjusted or price-standardized expenditures—that control for regional differences in prices—therefore more adequately reflect differences in utilization. *Utilization*, as used in the literature, is also commonly measured as counts (e.g., number of physician visits) or as rates of procedures in area-level analyses (e.g., number of surgeries performed per 1,000 beneficiaries). *Cost* is an ambiguous term and often implies a specific perspective (i.e., cost to whom, or spending from the perspective of a specific stakeholder), such as the cost to the patient, provider, or payer. Ultimately, these different terms, and the measures associated with them, seek to distinguish between patients receiving more versus less health care, irrespective of whether that measure is derived from the provider side (e.g., charges and accounting costs) or the payer side (e.g., Medicare expenditures or payments). As such, these measures are different representations of the intensity of healthcare services provided to and consumed by patients. This is the intention when *expenditure, utilization*, and *cost* are used in the remainder of this manuscript.

Measures of spending have been operationalized in a variety of ways: charges (the amount billed by providers to insurers);[125-128] accounting costs from providers' accounting systems;[8,129-150] care intensity indices (e.g., the Dartmouth Atlas' End-of-Life Expenditure Index) that reflect a relative amount of resource use in the healthcare production function;[4,5,151-161] and expenditures, measured as payments for healthcare services made by payers such as Medicare.[6,119,162-179] Charges reflect markups, and they likely overstate actual costs. Hussey and colleagues found that 39% of studies used accounting costs, 33% of studies measured expenditures, 21% used a care intensity index, and 7% used unadjusted charges.[23] Accounting costs were predominantly used in studies where the hospital (or nursing home) was the level of analysis. In contrast, care intensity indices were predominantly used in area-level analyses, although several such studies also used expenditures as measures of spending.[6,119,162,163]

Outcome Measures

A variety of dependent variables have been used as representations of quality or "benefit" derived from spending. These include measures of structure, process, and outcome (following Donabedian's model of healthcare quality),[180] patient experience (including satisfaction), access, or composites of these variables.[23] In general, the majority of studies used outcome measures of quality, with mortality being the most commonly used outcome.[4,8,125,127,128,130,132-134,136-142,147,151-153,158-160,162,165-167,170,179,181]

Mortality has been studied over a variety of timespans, from inpatient mortality,[152,167] to 30-day mortality,[166] to 1-year mortality, to 5-year mortality.[181] In a prominent series of papers, Fisher and colleagues (2003)[4,5] and Skinner and colleagues (2005)[181] examined the association between total Medicare spending and 1-year and 5-year mortality. In contrast, Kaestner and colleagues (2010)[127] focused on 30-day mortality in conjunction with inpatient Medicare spending only, in order to limit confounding by other sources of spending outside of the inpatient setting. Importantly, the studies by Fisher and colleagues examined HRR-level, rather than hospital-level, intensity of spending patterns. This is germane to the issue of spending because patients discharged from hospitals following inpatient stays often continue receiving services (such as outpatient care) within their HRRs.

Because the majority of studies have relied on administrative data, patient-reported outcomes and measures of physical, cognitive, and mental functioning (which are not commonly collected in such datasets) have received less attention in the literature. Such outcomes capture the elements of healthcare utilization that are not intended to alter the likelihood of survival, but rather to influence patients' health and functioning. As exceptions, Fisher and colleagues (2003) and Hadley and colleagues (2011)[179] used the Medicare Current Beneficiary Survey (MCBS) to examine patient-reported measures, such as access to care,[5] satisfaction with care, and functional status.[4] Functional status was measured using the Health Activities and Limitations Index (HALex).[182] Other measures of functional status, namely activities of daily living (ADLs), were limited to studies conducted in nursing homes where data on ADLs are routinely collected as part of the Minimum Data Set.[146,147,170] As one exception, Picone and colleagues (2003)[142] used binary indicators of ADL or instrumental ADL (IADL) limitations for a hospitalized patient population derived from the National Long-Term Care Survey linked to Medicare claims. The data used in this study are over 20 years old and do not provide current information about the effects of spending on ADL and IADL limitations in hospitalized patients. Taken together, there has been little emphasis in this literature on the non-mortality aspects of health status.

Methods to Address Confounding

Confounding by unobserved differences in patient characteristics—particularly health status—is a central challenge facing studies whose goal is to assess the consequences of regional variation in healthcare spending. To date, there are no experimental studies examining how differences in the intensity of spending or utilization affect outcomes. Randomized experiments could resolve confounding by observable and unobservable health status, but would be impractical due to the logistical and ethical challenges of randomizing patients to different levels of intensity and then ensuring that clinicians provide care at those level of intensity.

A central challenge of non-experimental research reported in this body of literature relates to addressing the endogeneity of spending measures. An endogenous explanatory variable is one that is correlated with the error term in a regression equation. Endogeneity can be caused by omitted variables, simultaneity, or measurement error.[183] In this context, bias due to omitted variables occurs when important variables—such as severity of illness measures—are unobserved in the data but are likely correlated with both spending/utilization and health outcomes. Simultaneity (sometimes described as reverse causality) occurs when the underlying causal pathway is as likely to be directed from health to spending as it is from spending to health. Jonathan Skinner has identified that area-level studies that entail simple comparisons of outcomes between higher-spending and lower-spending regions may fall victim to reverse causality because the sickest regions tend to spend more on health care.[181]

Confounding by severity of illness is difficult to overcome empirically. The geographic variation literature exhibits several approaches. These include multivariable regression analysis, which risk-adjusts only for the effects of observable health status (and the endogeneity problem is essentially ignored); natural experiments, whereby patients are assigned to exposures using a natural varying feature (not controlled by researchers) that is “as if” randomly assigned (such as regional intensity of end-of-life expenditures);[184] and instrumental variables estimation, which uses instrumental variables (observable variables that influence treatment or exposure but do not directly affect outcomes) to address unobservable health status.[185,186] A complicating concern with all of these approaches, but especially with multivariable regression, is that diagnostic practices—those that identify common comorbidities used for risk adjustment in regression models—also vary regionally. The implications of this were discussed in Part 1 of the literature review.

The majority of studies reviewed used multivariable regression adjustment.[23] Zhang and colleagues (2009)[168] used propensity score matching in an effort to reduce selection bias. Neither regression adjustment nor propensity scores ameliorate unobserved confounding.[187] Some studies used area- or region-level intensity measures as the key exposure variables. For example, Fisher and colleagues (2003)[4,5] used "exposure" to an area-level price-adjusted End-of-Life (EOL) Expenditure Index, which measured the intensity of care in the last 6 months of life (often called a "look-back" measure), as the main independent variable. This measure assumes that, across regions, all patients with an identical prognosis are equally ill, and so it intends to capture the part of regional variation in spending that is attributable to physician practice patterns rather than to regional differences in illness or price. The authors argue that this design naturally randomizes patients to different spending levels and therefore takes into account confounding by unobserved health status. The EOL Expenditure Index, and related measures (such as spending in the last 2 years, rather than in the last 6 months of life) have since been used in studies by Stukel and colleagues (2012) studying the association of hospital spending intensity with mortality and readmission rates in Ontario hospitals,[188] O'Hare and colleagues (2010)[189] in examining regional variation in end-stage renal disease treatment practices, and Hadley and colleagues (2011) in studying the relationship between cumulative Medicare spending and beneficiaries' mortality and functional status.[179] However, Bach and colleagues

(2004 and 2010)[190,191] have criticized such look-back measures as misleading and argued that differences in subject selection and time period can bias results about the care provided to decedents. Other area-level measures have also been used. Stukel and colleagues (2005)[192] used area-level cardiac catheterization or beta-blocker prescription rates as primary exposures. In a distinct natural experiment, Doyle and colleagues (2007)[152] assessed the outcomes of patients who received care in healthcare systems not near their home; in this case, the patients were visitors to Florida who experienced heart attacks and were hospitalized; the study used multivariable regression as well as an alternative analysis using instrumental variables.
A minority of studies used instrumental variables estimation as the analytic approach.[23] An instrumental variable, or "instrument," is an auxiliary variable that generates variation in the endogenous variable of interest but has no direct effect on the outcome variable(s) and is uncorrelated with unobserved confounders. The goal of this method is to determine how much of the variation in the endogenous variable is induced by the instrument itself (the so-called exogenous variation) and how that affects the outcome variable. The induced variation serves to identify the desired estimate.[186] The concept of identification—in an econometric sense—refers to the possibility of estimating a causal effect by writing a parameter in terms of population moments that can be estimated using a sample of data generated by a data-generating process.[193,194]

Examples of instruments used in these studies included care intensity indexes[127,152,179] (such as the EOL Expenditure Index), physician visits or intensive care unit (ICU) days in the last 6 months of life,[181] ambulance referral patterns,[150] hospital demand,[143] the Medicare Wage Index and general overhead costs per day at the hospital level,[195] and outcome values from previous years.[130] Instrumental variable estimation, unlike ordinary least squares (OLS), can yield consistent estimates of the effect of spending on health outcomes. Consistent estimates are vital for informed policymaking. Because spending is endogenous and likely positively correlated with unobserved severity of illness measures, the coefficient on spending would be biased toward showing that more spending worsens outcomes if estimated by OLS.[179,193] In general, the literature supports the view that studies using more rigorous research designs, such as instrumental variables analysis, have found that additional healthcare spending is associated with improved outcomes.[150,179,196-198]

Instrumental Variables Estimation

Because of its importance within this literature, additional discussion about instrumental variables estimation methodology follows. Finding suitable candidate instruments is a key challenge of using instrumental variables estimation. An instrument should have two characteristics: validity and relevance.[193] Instrument validity is based on whether the instrument is exogenous (i.e., uncorrelated with unobserved severity of illness) and has no direct effect on health outcomes (the exclusion restriction).[186] For example, Hadley and colleagues (2011)[179] used the EOL Expenditure Index as an instrumental variable, an example of a measure of regional treatment intensity.

Suitability of measures of regional spending or treatment intensity can be justified on the following grounds. First, evidence indicates that significant regional variation in spending and utilization persist even after accounting for a comprehensive set of patient characteristics and health status measures, suggesting that this variation reflects factors unrelated to patient health, such as supply-side practice patterns.[37] Kelley and colleagues (2011)[199] found that marked variation in EOL spending remained even after accounting for numerous risk-adjusters derived from Health and Retirement Study survey responses and Medicare claims. Second, physician factors and practice patterns are strongly associated with differences in spending across regions, supporting the viability of these candidate instruments. Sirovich and colleagues (2008)[96] surveyed physicians in HRRs around the U.S. by using identical clinical vignettes and asking participants about how they would treat the hypothetical patients. Physicians in higher-spending regions were more likely to recommend discretionary services (e.g., tests of unproven benefit) and also to schedule more frequent return visits when compared to physicians in lower-spending regions. Third, Fisher and colleagues (2003)[4,5] demonstrated that the average baseline health of patients in different HRRs was similar across quintiles of EOL treatment intensity. Stukel and colleagues (2005)[192] similarly demonstrated that baseline heart attack severity was similar across regions characterized by widely varying treatment intensities. Both studies suggest that regional measures of treatment intensity do not directly affect health outcomes.[200] Instrument validity is often justified on a theoretical basis, and many papers using instrumental variable estimation devote considerable explanation to justifying the validity of the chosen instruments. If valid instruments are not available, then instrumental variables estimation is not an appropriate analytic method.

Instrument relevance is based on whether or not an instrument explains variation in the endogenous variable (whether it is positively or negative correlated with it). Unlike validity, instrument relevance can (and should) be tested using the sample of data.[193] Commonly used means of assessing instrument relevance include the partial R^2 value, Cragg-Donald F statistic, and Kleibergen-Paap Wald statistic. The partial R^2 is the R^2 between the endogenous variable and the instrument after controlling for other covariates in the model. The Cragg-Donald F statistic and Kleibergen-Paap Wald statistic are tests for weak instruments; the Kleibergen-Paap Wald test is a robust analog of Cragg-Donald.

A commonly used rule of thumb proposed by Staiger and Stock (1997)[201] is that an F statistic below 10 suggests the possibility of a weak instrument problem. Stock and Yogo (2005)[202] offer an alternative approach based on the Cragg-Donald F statistic in the case of a single endogenous variable (with the null hypothesis that the instruments are weak) and critical values based on size distortion of the 5% Wald test. In the case of clustered standard errors, Baum and colleagues (2007)[203] suggest either applying these critical values to the Kleibergen-Paap test with caution or using the original Staiger and Stock (1997) rule-of-thumb that the F statistic should be 10 or more.

Findings from the Literature

Findings from studies comprising the growing empirical literature examining the spending-outcome relationship are found to present a mixed picture: some studies identify a positive association between spending and outcomes, while others found a negative or mixed association, and yet others found no association. In part, this equivocal evidence stems from the marked heterogeneity in study designs, data sources, and measures used. In general, studies finding a positive association between spending and outcomes (most commonly, mortality) had magnitudes conveying minimal clinical significance.[23]

<u>Negative, Null, and Mixed Findings</u>

Among those finding negative associations, Baicker and colleagues (2004)[119] found a negative association between per capita Medicare spending and overall state quality ranking; and, on individual quality measures, they found either significant and negative associations or nonsignificant associations. Comparing HRRs in the highest- to the lowest-spending quintile based on the EOL Expenditure Index, Fisher and colleagues (2003)[4,5] found that patients with acute myocardial infarction (AMI) receive a lower quality of care (were less likely to receive acute reperfusion, aspirin at admission or discharge, and ACE inhibitors at discharge). Higher-spending regions were also less likely to provide flu immunizations, pneumonia immunizations, and Pap smears.[5] They found that risk of death was worse for colorectal cancer and AMI patients and not significantly different for hip fracture patients in regions with higher EOL spending. But they found that functional status was not significantly different between the highest- and lowest-spending regions.[4] In a study specific to colorectal cancer that also used the EOL Expenditure Index, Landrum and colleagues (2008)[153] found that greater spending intensity was significantly associated with increased *non-cancer* mortality but not significantly associated with all-cause or cancer-specific mortality. Looking at a process measure of chemotherapy receipt, they found that higher-spending regions had a greater propensity to use recommended care, but also nonrecommended and discretionary care.

With regard to patient perceptions of care, Fowler and colleagues (2008)[163] failed to find significant differences in the patient having a perceived unmet need for tests or treatments, or in a patient's perceived quality of care (ambulatory care and overall care) when comparing the highest- and lowest-spending regions. In contrast, Fisher (2003)[4] and Wennberg (2009)[155] found that overall patient satisfaction or experience was negatively associated with care intensity. Similarly, Yasaitis and colleagues (2009)[156] found that compared with patients in lower-intensity regions, those receiving health care in higher-intensity regions reported lower efficiency, satisfaction, overall quality of care, and patient centeredness, and lower ratings of physician-patient communication.

For the assumption and observed negative associations that more spending undermines outcomes, possible mechanisms are suggested by several studies showing that regions with higher utilization rates perform certain interventions in patient populations with lower underlying medical need—a signal of possible overuse.[151] The burgeoning overuse literature[204] is related to, though distinct from, the practice variation literature. Variation in utilization may imply, but does not prove, overuse; especially if differences in utilization are related to "appropriate" sources of variation, such as patient preferences.[205] (Note: Patient preferences, and other determinants of variation in spending, were discussed in Part 1 of this literature review.) A systematic review by Keyhani and colleagues (2012)[206] found limited evidence that inappropriate use of procedures explains geographic variation in spending and utilization. By using a distinct conceptual framework focused on identifying *systematic* overuse of health care (i.e., utilization that is expected to have no health benefit and is likely to be the result of physician factors), Segal and colleagues (2014)[207] developed an index of overuse at the HRR level and emphasized that overuse is a phenomenon that is distinct from that of utilization. They found that regional overuse was correlated with overall utilization and neither positively or negatively correlated with mortality.[208]

Positive Findings

At a hospital level, numerous studies have found positive associations between higher costs and mortality. In assessing inpatient mortality, Romley and colleagues (2011)[167] found that higher-spending hospitals had lower risk-adjusted inpatient mortality among patients hospitalized with diagnoses of acute myocardial infarction, congestive heart failure (CHF), stroke, gastrointestinal hemorrhage, hip fracture, and pneumonia. Similarly, Kaestner and Silber (2010)[127] used instrumental variables analysis and found that among Medicare patients hospitalized for general, orthopedic, and vascular surgery or for CHF, stroke, and gastrointestinal bleeding, a 10% increase in hospital spending (charges) was associated with a 3.1% to 11.3% decrease in 30-day mortality (depending on the disease). In a pair of studies analyzing data for hospitals in California, Mukamel and colleagues (2001 and 2002) found that additional spending per discharge was associated with lower 30-day mortality per 100 discharges[141] and that hospitals above the 50th percentile for cost per admission had lower mortality rates.[140] Conversely, Schreyogg and colleagues (2010)[195] analyzed spending during index hospitalizations for acute myocardial infarctions, using instrumental variables analysis; they found that lower costs were associated with a 0.63% increase in the risk of dying and a 1.24% increase in the risk of readmission. Using longer intervals for outcome assessment, Doyle and colleagues (2012)[150] and Picone and colleagues (2003) found that increases in the cost of a hospital stay were associated with lower 1-year[150] and 2-year[142] mortality rates.

Barnato and colleagues (2010)[158] also assessed hospitals but instead used a revised EOL Expenditure Index (different from the one described previously that was developed by researchers at Dartmouth) rather than hospital cost, and they found that admission to higher EOL treatment intensity hospitals was associated with lower 30-day mortality. Stukel and colleagues (2012)[188] also used a variant of the EOL Expenditure Index as the "primary exposure" to treatment intensity for adults hospitalized in Ontario hospitals for AMI, CHF, hip fracture, or colon cancer. They found reductions in mortality, readmissions, and cardiac event rates for patients admitted to higher-spending intensity hospitals. Using a similar measure of EOL intensity but at the county-level, Doyle (2007)[152] found that a 10% increase in county EOL intensity was associated with a 0.3–percentage point decrease in mortality.
Several other studies examined nonmortality outcomes. Carey and Burgess (1999)[130] found improved outpatient follow-up rates 30 days after inpatient discharge (process measure) in hospitals with higher costs. Hadley and colleagues (2011)[179] evaluated the HALex score using MCBS data from 1992-2002 and instrumental variables estimation and found that more spending was associated with improved functional status. At a state level, Richard Cooper (2009)[6] found that greater total per-capita healthcare spending in 2004 was associated with better state-level quality (based on Jencks quality rankings of health system performance).

When viewed collectively, this literature presents an equivocal picture with respect to the question of whether greater intensity of spending or utilization—either at a hospital or in a region—is associated with improved quality of care and patient outcomes. The studies reviewed are heterogeneous in their study designs, employing a diversity of spending and outcome measures, patient populations, and empirical strategies. Two important observations can be drawn from this part of the review, and these observations can inform future research: (1) unobserved health status should be dealt with through more rigorous study designs and methods, and (2) there is relative dearth of information about the effect that spending has on nonmortality outcomes, such as measures of physical, cognitive, and mental health and functioning.

CONCEPTUAL FRAMEWORK

A conceptual framework was developed based on the preceding literature review that integrates determinants of regional variation in utilization and identifies key variables. No single unifying conceptual framework underlies research in medical practice variation.[13] This literature emanates from a diversity of disciplines: epidemiology, beginning with John Wennberg's pioneering work on small area variation;[2] economics, which has sought to parse out supply and demand factors contributing to geographic variation;[22] and sociology, which has emphasized the social context influencing physicians' decision-making behavior as it relates to practice variation.[94,102,103,209] These perspectives are not mutually exclusive, but have employed different theories and methodologies to study practice variation.

Acknowledging these differing traditions, Ronald Andersen's widely known behavioral model of health services utilization,[210] which has an extensive history in health services research for studying utilization, was adapted. Early versions of this model focused on measures of access, emphasizing utilization as the endpoint of interest; later versions focused on health outcomes as the endpoints of interest. Andersen's conceptual model accommodates determinants of utilization that exist at multiple levels (e.g., individual and regional factors). One criticism of Andersen's original model was that it was biased towards assuming that increased utilization was always better.[211] Andersen has defended the model, believing it is non-normative regarding utilization,[210] and the adaptation of Andersen's model here makes no such normative judgments.

Figure 2.1 depicts a modified version of Andersen's model. The grey boxes represent the important conceptual areas originally described by Andersen, and the white boxes beneath indicate the key factors pertinent to the study of variation in spending or utilization. For brevity, not every demand-side or supply-side factor discussed in Part 1 of the literature review is depicted in Figure 2.1; the bullets are examples of some of the most important factors.

In this model, the environment comprises elements of the healthcare system, such as region- and hospital-level factors. These interact with characteristics of the population, including the population's predisposing characteristics, such as its sociodemographics, health behaviors, and care preferences; enabling resources (factors that support or hinder use of health care), such as access to care, health insurance, and income or wealth; and its need, which has both perceived and evaluative components. Perceived need is how an individual interprets his or her health status and makes decisions to seek professional assistance, whereas evaluative need reflects physicians' professional judgment and the type and amount of care needed by the patient.[210] Therefore, "need" can refer to the onset of an unambiguous disease necessitating treatment or intervention (e.g., heart attack)—what Andersen calls the "most immediate cause of health service use"[212] or "need" can refer to an interpretation based on the patient or physician's perception or judgment of the value of intervention. Collectively, these environment and population characteristics influence the use of healthcare services, which Andersen describes as "health behavior" (or, more fittingly, healthcare behavior). In this manuscript, use of health services has been described as being represented by spending (e.g., spending in the year following hospitalization) or utilization (e.g., number of inpatient admissions in a given time period). Use of health services ultimately impacts outcomes and patients' final health status.

Application of Conceptual Framework to Studying Geographic Variation

The conceptual framework can be applied to studying the effects of variation in Medicare spending on a myriad of outcome measures. Analyses can harness geographic variation in practice patterns (***Healthcare System*** characteristics), such as measures of regional intensity at the EOL, as a source of possible instrumental variables to estimate causal effects while controlling for ***Predisposing Characteristics*** and ***Enabling Resources***. The dashed line extending backward from ***Outcomes*** to ***Use of Health Services*** illustrates the possible endogeneity concern discussed previously (here depicted as a form of reverse causality).

In Andersen's model, hospital services received in response to more serious diseases and conditions would be predominantly explained by the categories ***Need*** and ***Predisposing Characteristics*** (e.g., baseline health status and demographics).[210] Researchers at Dartmouth and others have often identified study populations on the basis of medical need by selecting subgroups of patients hospitalized with acute episodes of specific diseases (e.g., AMI, hip fracture, stroke, gastrointestinal bleeding) because these exhibit minimal variation in their hospitalization rates across regions, suggesting that patients with incident cases likely have similar severity of illness regardless of their geographic location.[121] ***Predisposing Characteristics*** and ***Enabling Resources*** (e.g., wealth) are important control variables and represent demand-side factors discussed previously.

The conceptual framework also provides an opportunity to analyze additional understudied determinants of spending and utilization. Much of the variation in spending across regions remains unexplained.[36] For example, health behaviors and modifiable risk factors—smoking, obesity, alcohol consumption, and physical activity—rather than practice styles of varying intensity may be important causal factors explaining regional variation in utilization.[22] In the conceptual model, health behaviors are a type of predisposing characteristic emanating from beliefs and values concerning health and illnesses.[213] In Chapter 3 (Aim 2), decomposition techniques are used to quantify the contribution of health behaviors and patient characteristics to explain regional variation in the spending and utilization (***Use of Health Services***). Implementation of this method requires using Dartmouth Atlas measures of regional treatment intensity to divide respondents into mutually exclusive groups of higher-versus-lower regional intensity (***Healthcare System*** characteristics) based on where they live and then estimating separate regressions for each group. This decomposition offers a way to differentiate between differences in utilization that derive from observable differences in group ***Predisposing Characteristics*** (specifically, health behaviors and modifiable risk factors) versus differences owing to unexplained factors, which plausibly include provider practice patterns in how the groups are treated.

DISCUSSION

This review has emphasized the determinants and consequences of geographic variation in healthcare spending and utilization by bringing these diverse literatures together and adapting a well-known conceptual framework to describe how these demand-side and supply-side factors contribute to healthcare spending and outcomes. When viewed collectively, the body of evidence appraised herein suggests several important directions and considerations for future research. As identified in Part 1 of the review, health behaviors and modifiable risk factors as possible determinants of variation in spending have been understudied because data on smoking, alcohol consumption, diet, and physical activity or exercise are rarely collected in administrative datasets. Similarly, a hallmark of the literature on the consequences of geographic variation in spending has been its predominant use of administrative data, such as Medicare claims, and its focus on mortality and survival as the primary representations of health. Consequently, studies of non-mortality outcomes (capturing, health status, functioning, and health-related quality of life) are limited in number. Yet because much of health care aims to improve health-related quality of life, and not solely mortality or survival, assessing the effects of spending on a comprehensive set of outcomes is an important direction for future studies. Conducting such studies requires datasets with broad geographic coverage that also collect data on such outcomes, a feature jointly offered by few datasets.

In addition, more rigorous research designs can improve inference about whether additional healthcare spending is causally related to outcomes. While the majority of such studies examined in this manuscript used standard multiple regression approaches and were likely limited by omitted variable bias, there is a precedent in the health economics literature for using instrumental variables analysis—an approach that is adopted in Chapter 4 (Aim 3). Instrumental variables analysis can provide policymakers with more accurate information about the effects of healthcare spending on patient outcomes.

This review also suggests possible challenges to anticipate when planning new studies. First, as has been demonstrated with mortality, measurable benefit from treatment intensity may wane with time,[158] and this may be true for other outcomes as well, complicating the ability to detect an effect for outcomes measured long after the initial hospitalization. Second, the assessment of outcome measures should ideally be related to the level of analysis and the specification of spending measures. For example, it may be difficult to detect an effect of *inpatient* utilization/spending following an index hospitalization on long-term mortality due to confounding by other sources of utilization that occur between the time utilization is measured and mortality is assessed (e.g., outpatient and post-acute utilization). Such “history effects,” as noted in the lexicon of Shadish, Cook, and Campbell (2001),[214] can be minimized through study design considerations. Region-level analyses might benefit from analyzing longer-term outcomes and aggregate spending whereas hospital-level studies might benefit from using shorter-term outcomes and inpatient spending, which are more tightly coupled with the effects of the hospital.

Third, studies examining outcomes other than mortality, such as health-related quality of life, face the problem of death as a competing risk. Truncation due to death is an important threat to validity, because survivors with observed outcomes are not a random sample of the study population. There are a variety of approaches for handling truncation due to death. One straightforward approach is to use composite outcomes that combine information about survival and the health-related quality of life measure.[215] Alternatively, sensitivity analysis procedures formalized within a principal stratification framework can be used to draw inferences about survivor average causal effects.[216] Principal stratification has several advantages over other approaches, such as selection models, because patient-reported outcomes are undefined when the outcomes are not observed.[215,217]

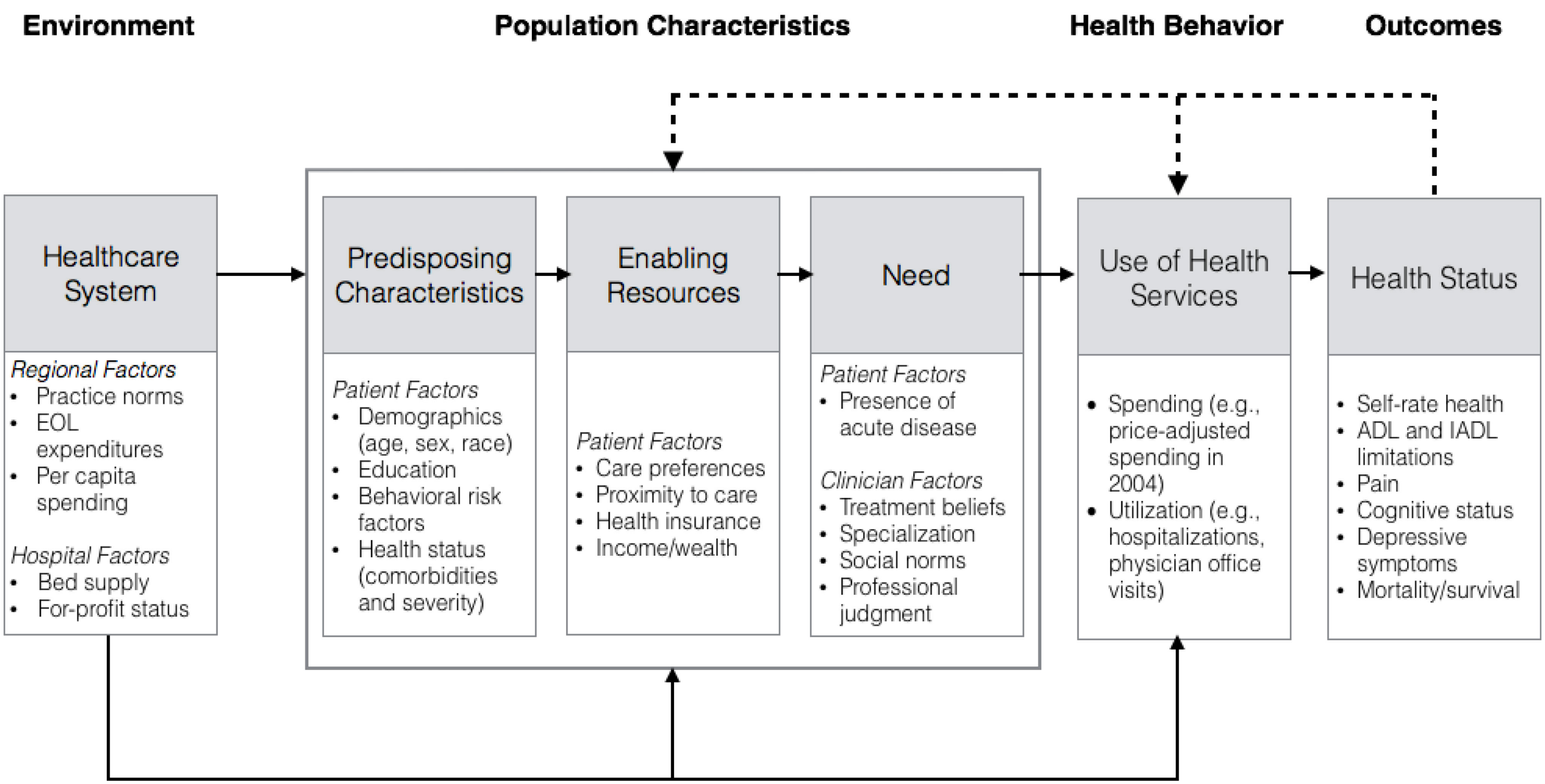
Fourth, care must be taken when using comorbidities derived from claims data due to the risk of observational intensity bias. Nevertheless, confounding by severity of illness is a central challenge facing studies assessing the consequences of geographic variation. Concerns inherent in risk adjustment can be addressed by (1) identifying more homogenous disease-specific subgroups; (2) considering the inclusion of additional health status characteristics derived directly from patients—such as smoking status, alcohol consumption, and body mass index—that are more likely to be free from observational intensity bias or variation in physician diagnostic practices; and (3) using instrumental variables estimation, which can mitigate biases from unmeasured confounders, as described previously. In its report, the IOM advocated for health status measures derived from claims data to be supplemented by behavioral and clinical data for Medicare and commercially insured beneficiaries.[3]

In light of this discussion, this dissertation addresses important gaps in current knowledge about the determinants and consequences of geographic variation in spending for older Americans.

First, the dissertation will analyze a comprehensive set of outcome measures related to multiple dimensions of health and functioning—activities of daily living, pain, cognition, and depressive symptoms—that are essential to the quality of life of the aging population, but are understudied in the geographic variation literature. Second, this dissertation extends previous research by using econometric methods that facilitate causal inference for questions salient to policymakers. Regression decomposition techniques quantify the contribution of health behaviors, an understudied set of determinants of utilization, by explaining variation in spending and utilization across regions. Instrumental variables estimation is used to estimate causal effects of healthcare spending/utilization on outcomes, an improvement on existing observational research in this literature. These contributions are made possible by leveraging a novel dataset—the Health and Retirement Study (HRS) linked to Medicare claims—that combines broad geographic coverage with data on utilization (via the claims) and patient-reported health outcomes and behaviors (via the survey), rendering it ideally suited to the aims of this dissertation and for advancing an understanding of the determinants and consequences of geographic variation in healthcare spending.[218,219]

FIGURES

Figure 2.1. Conceptual Model



Abbreviations: EOL, end of life; ADL, activity of daily living; IADL, instrumental activity of daily living

CHAPTER THREE (MANUSCRIPT #2)

RELATIONSHIP BETWEEN GEOGRAPHIC VARIATION IN MEDICARE SPENDING AND BENEFICIARIES' BEHAVIORAL RISK FACTORS

by

Kurt Richard Herzer

ABSTRACT

Importance: Perennial debate about the pronounced and persistent variation in Medicare spending across regions of the United States centers on whether variation is due to inefficiencies in providers' practice patterns or to beneficiaries' characteristics. In particular, little is known about how beneficiaries' behavioral risk factors contribute to geographic variation in healthcare spending.

Objective: To examine beneficiaries' behavioral risk factors and their relationship to regional variation in Medicare spending.

Design: Secondary analysis of data from the nationally representative Health and Retirement Study linked to Medicare claims and regional spending characteristics from *The Dartmouth Atlas of Health Care*. Regression-based decomposition analysis was used to assess the contributions of beneficiaries' characteristics (sociodemographics, behavioral risk factors, and health status) to Medicare spending across higher- versus lower-spending regions.

Setting: Hospital referral regions in the United States.

Participants: Medicare beneficiaries, aged 65 or older, in fee-for-service Medicare in 2004.

Exposures: Smoking status, alcohol consumption, body mass index, and physical activity.

Main Outcomes and Measures: Medicare Part A and B spending for 2004-2006 (price-adjusted, expressed in 2006 U.S. dollars). The percent of regional variation in spending explained by different beneficiary characteristics was calculated.

Results: Among 8,476 Medicare beneficiaries, individual characteristics explained 17% of regional differences in price-adjusted Medicare spending. Behavioral risk factors collectively explained 7% of the $5,718 difference in spending between higher- and lower-spending regions.

Conclusions: Although the majority of regional variation in Medicare spending was not explained by beneficiary characteristics, policies designed to mitigate behavioral risk factors among Medicare beneficiaries in higher-spending regions may provide an opportunity to

modestly reduce geographic variation in spending. Additional research is needed to better elucidate non-beneficiary characteristics that may be particularly amenable to policy interventions.

INTRODUCTION

For decades, policymakers have been interested in longstanding variation in Medicare spending across regions of the United States.[220] The perennial debate—centered on which factors contribute to this geographic variation—has been framed by divergent perspectives about the relative importance of beneficiary-versus-provider characteristics.[35-37] While prior research has demonstrated that beneficiary characteristics such as health status account for a small-to-moderate proportion of the variation in healthcare spending across regions,[37,199] other important factors such as beneficiaries' health behaviors and modifiable risk factors have not been adequately studied as possible drivers of regional variation in spending.[18]

Behavioral risk factors, specifically smoking, alcohol consumption, obesity, and limited physical activity, are prevalent in the aging American population,[14] vary geographically,[15,16] and represent a major source of growth in healthcare spending in the preceding decades.[17] Policymakers need better information about whether targeting the modifiable antecedent causes of disease—beneficiaries' lifestyle factors—could reduce spending in higher-spending regions.

To address this gap, we linked Medicare beneficiaries' self-reported data on health behaviors and modifiable risk factors from the Health and Retirement Study (HRS) to utilization data contained in Medicare claims. The objective of this study was to examine whether behavioral risk factors contribute to explaining the marked variation in Medicare spending across regions of the United States.

METHODS

Data Sources and Sample

The primary source of data was the HRS, a large, nationally representative, prospective cohort study of older Americans that has been used in prior research to examine geographic variation in Medicare spending.[199,221-223] We linked survey measures of sociodemographic characteristics, behavioral risk factors, and health and functioning from the HRS to respondents' Medicare claims. We studied respondents who participated in the 2004 wave of the HRS, in which, 86% of eligible respondents consented to release their Medicare claims for research purposes.

Because we wanted to observe beneficiaries' actual Medicare spending, we could only study those who were enrolled in traditional (fee-for-service) Medicare since claims data are not available for those who select a managed care plan through the Medicare Advantage option. In 2004, Medicare Advantage enrollment was at an historic low nationally (13%) during the 1999-2015 period (and approximately 17% among HRS-Medicare linked respondents in our sample).[224] Use of 2004 as the base year served to reduce sample selection bias from those beneficiaries selecting into Medicare Advantage plans who might differ in health status, health behaviors, or in unmeasured ways.

We restricted the study population to respondents who were 65 years or older (N = 10,393), enrolled in fee-for-service Medicare (N = 8,532) during 2004, and for whom employer-sponsored insurance was not the primary payer (N = 8,476). To prevent reductions in sample size, we did not exclude respondents who switched to Medicare Advantage plans after 2004 or who did not have full enrollment in both Medicare Parts A and B during the study period. Instead, we included variables to model these coverage characteristics and conducted a sensitivity analysis to explore the effect of excluding these respondents.

Behavioral Risk Factors

We assessed respondents' smoking status, alcohol consumption, body mass index (BMI), and physical activity because these are commonly studied behavioral risk factors and improvements to these factors could reduce healthcare costs and premature deaths from cardiovascular disease, cancer, and other causes.[225] Based on respondents' HRS survey responses, we divided smoking status into three categories: nonsmoker, former smoker, or current smoker. Alcohol consumption was based on respondents' use within a 3-month reference period. It was categorized using a quantity-frequency approach similar to that used in prior studies: abstainers (no alcohol use), light drinkers (1 to 3 drinks per week), moderate drinkers (4 to 7 drinks per week for women; 4 to 14 drinks per week for men), heavy drinkers (8 to 34 drinks per week for women; 15 to 34 drinks per week for men), and very heavy drinkers (≥35 drinks per week for both sexes).[226] BMI was calculated using respondents' self-reported height and weight and classified into the following categories based on standard Centers for Disease Control and Prevention (CDC) criteria: underweight (BMI below 18.5), normal weight (18.5 to 24.9), overweight (25.0 to 29.9), or obese (30 or above).[227] We also assessed respondents' physical activity—whether they engaged in light, moderate, and vigorous physical activity at least weekly (these categories were not mutually exclusive). Light physical activity included activities such as vacuuming and doing laundry and home repairs. Moderate exertion referred to activities such as gardening, walking at a moderate pace, and performing stretching exercises. Vigorous exertion included activities such as running or jogging, swimming, cycling, and gym workouts.

Medicare Spending and Utilization

Our primary outcome measure was total Medicare spending for these respondents over the 3-year period from 2004 to 2006. This cumulative measure included spending for care in a variety of settings (inpatient, skilled nursing facilities, outpatient, physician office, home health, and hospice), as well as spending for durable medical equipment. For each year, we adjusted spending for regional differences in price using the ratio of price-standardized to price-unstandardized Medicare Part A and B spending within respondents' hospital referral regions. Hospital referral regions (HRR) are local markets wherein beneficiaries receive the preponderance of their health care.[220] Price adjustment reduces the effects of the varying labor, capital, and overhead costs encountered across regions and serves to more accurately reflect the quantity of care utilized.[124] Total Medicare spending from 2004 to 2006 was expressed in 2006 U.S. dollars based on the medical care component of the Consumer Price Index. As secondary outcome measures, we examined the number of inpatient admissions and outpatient or physician office visits incurred by respondents between 2004 and 2006.

Regional Spending Level

We divided respondents into mutually exclusive groups of higher- and lower-spending regions based on whether their HRR was in the top or bottom half, respectively, of the national distribution of price-adjusted Medicare Part A and B spending. These regional Medicare spending data were obtained from *The Dartmouth Atlas of Health Care*.[19]

Statistical Analysis

Our empirical analysis centered on determining how much of the observed variation in Medicare spending across U.S. regions could be explained by beneficiaries' behavioral risk factors. To this end, we used a regression-based decomposition analysis originally developed by labor economists Blinder and Oaxaca to study wage gaps between men and women,[228,229] and which has been applied in health policy and geographic variation research.[230-232] We used this method to quantify how much of the difference in Medicare spending between higher- and lower-spending regions could be explained by differences in beneficiaries' observable characteristics (particularly, their behavioral risk factors) versus differences in the way that providers in those different regions would treat the same beneficiary.[231,233,234]

We used multivariate regression models to examine the relationship between Medicare spending and beneficiaries' behavioral risk factors. To model positively skewed Medicare spending, we estimated generalized linear models with a gamma distribution and log link.[235,236]

To model utilization outcomes—inpatient admissions and outpatient or physician office visits—we fit generalized linear models with a negative binomial distribution and log link. Standard errors were clustered by region.[237] To determine the contribution of behavioral risk factors (and other beneficiary characteristics) to explaining regional variation in spending, we compared the regression-adjusted Medicare spending for beneficiaries in lower-spending regions with the spending that would have been predicted if they had the distribution of sociodemographic characteristics, behavioral risk factors, health and functioning, and insurance coverage of beneficiaries in higher-spending regions.[234,238] We then calculated the proportion of the regional difference in Medicare spending attributable to beneficiaries' behavioral risk factors and other characteristics.

In addition to behavioral risk factors, our regressions also accounted for the following characteristics that may vary across regions: sociodemographic characteristics (including age, sex, race/ethnicity, educational achievement, household wealth, poverty status, marital status, household size, and urban status), health and functional status (including self-rated health, proxy respondent status [as an indirect measure of health], death between 2004-2006, limitations in activities of daily living (ADLs) and instrumental activities of daily living (IADLs), and cognitive status), and insurance coverage (including whether respondents had any enrollment [at least 1 month] in a Medicare Advantage plan in 2005 or 2006, any discontinuous enrollment in both Medicare Parts A and B [at least 1 month] between 2004 and 2006, and supplementary private insurance). In the main analyses, we did not control for diseases or conditions in order to avoid introducing a bias associated with regional variation in physicians' diagnostic practices and observational intensities.[42]

We conducted numerous sensitivity analyses to assess the robustness of our results. First, we repeated our analyses using different outcome measures, including a longer 5-year interval (spending and utilization between 2004 and 2008) and using 2006 as the base year (spending and utilization between 2006 and 2008) for HRS respondents who participated in the 2006 survey wave. Second, we examined the effect of including disease and condition variables, which are plausibly associated with regional variation in diagnostic practices,[42] similar to prior studies. Third, we repeated our analyses after correcting the BMI variable for possible self-reporting bias.[239] Fourth, we assessed the sensitivity of our results to alterations in the composition of the study sample by repeating our analyses after doing the following: excluding veterans (for whom the Department of Veterans Affairs may be the primary payer), excluding respondents who did not have continuous Part A and B coverage over the 3-year observation period, and excluding respondents who died after 2004.

Fifth, to address the concern that respondents in higher- versus lower-spending regions may self-select differently into Medicare Advantage during the observation period, we compared characteristics of beneficiaries “lost” to Medicare Advantage to fee-for-service beneficiaries remaining in the sample. We examined whether beneficiaries using Medicare Advantage differed on baseline characteristics in higher- versus lower-spending regions. We also implemented selection models to correct for possible sample selection bias owing to Medicare Advantage enrollment. Lastly, we examined versions of the standard linear Blinder-Oaxaca decomposition.[229] Details of all sensitivity analyses are provided in the Appendix.

We used Stata 14.0 (StataCorp, College Station, Texas) for all analyses. Reported P values are two-sided; a P value of 0.05 or less designated statistical significance. This study was approved by the Internal Review Board of The Johns Hopkins Bloomberg School of Public Health.

RESULTS

Characteristics of the Study Sample

In 2004, 11,399 of 13,175 HRS respondents with linked Medicare claims (86.5%) were interviewed. Of these, we excluded 1,006 respondents who were under age 65, 1,861 who were enrolled in Medicare Advantage plans during 2004, and 56 who reported having primary health insurance coverage through an employer. The remaining sample included 8,476 beneficiaries.

The mean age of all beneficiaries was 75.2 years; 58% were female, 21% were nonwhite, 10% were below the federal poverty threshold, and 69% lived in urban areas. Beneficiaries living in lower-spending hospital referral regions (median regional spending in 2004 was $6,680; range, $5,112 to $7,324) varied considerably on most of the observable characteristics from beneficiaries in higher-spending regions (median regional spending in 2004 was $7,986; range, $7,330 to $11,609) (Table 3.1). In particular, beneficiaries in higher-spending regions were more likely to be nonwhite, have lower educational achievement and wealth, and to live in urban areas. These beneficiaries were also more likely to abstain from drinking alcohol (64% versus 58%; $P<0.001$) and were less likely to engage in weekly moderate or light physical activity (59% versus 63%; $P<0.001$, and 72% versus 76%; $P<0.001$, respectively). These same beneficiaries generally reported a lower health status and were less likely to have supplemental private insurance. Beneficiaries in higher- versus lower-spending regions did not differ in their likelihood of having any Medicare Advantage enrollment in 2005 or 2006.

On average, total Medicare spending for beneficiaries between 2004 and 2006 was $27,759 with 1.1 inpatient admissions, and 32.0 outpatient facility or physician office visits (Table 3.1). In unadjusted comparisons, beneficiaries in higher-spending regions had $5,545 higher average Medicare spending compared to beneficiaries in lower-spending regions ($P<0.001$).

Factors Explaining Regional Variation in Spending and Utilization

Figure 3.1 demonstrates the effects of sequentially controlling for additional sets of variables in the regression model on the difference in Medicare spending between higher- and lower-spending regions. Based on the unadjusted model, which accounts only for regional differences in price, Medicare spending was 22% higher comparing higher- to lower-spending regions. The absolute difference in Medicare spending between higher- and lower-spending regions decreased by 25% and 30% when sociodemographic characteristics and then behavioral risk factors were added, respectively.

Overall, beneficiary characteristics explained 17% of the regional differences in Medicare spending (Table 3.2). Among the observable characteristics, behavioral risk factors and health status were important contributors to the lower spending for beneficiaries in lower-spending regions, explaining 7% and 14%, respectively, of the $5,718 difference in predicted spending between beneficiaries in higher- and lower-spending regions. While these characteristics contributed toward increasing the difference in spending between higher-versus-lower-spending regions, other characteristics decreased the difference. For example, insurance coverage characteristics—specifically supplementary private insurance coverage, which beneficiaries in lower-spending regions were more likely to possess—reduced the regional difference in spending by 6%. Similar relationships were observed between beneficiary characteristics and inpatient admissions, whereas nearly all of the regional difference in outpatient and physician office visits was not explained by beneficiary characteristics.

Figure 3.2 shows the predicted Medicare spending for beneficiaries in lower- and higher-spending regions. If beneficiaries in lower-spending regions had the same distribution of characteristics (e.g., behavioral risk factors and health status factors) as beneficiaries in higher-spending regions, their Medicare spending would be $956 higher; 83% of the regional difference in Medicare spending cannot be attributed to differences in beneficiaries' observable characteristics.

Across most characteristics, beneficiaries in higher-spending regions incurred significantly higher Medicare spending between 2004 and 2006 than beneficiaries in lower-spending regions (Figure 3.3; Appendix Table S3.1). For example, beneficiaries in higher-spending regions who were overweight or former smokers had $8,024 (95% CI, $4,675 to 11,372) and $7,637 (95% CI, $3,937 to $11,337) higher spending, respectively, than beneficiaries who were overweight or former smokers in lower-spending regions. Models estimated separately for beneficiaries in lower- and higher-spending regions are presented in Appendix Tables S3.2 and S3.3, respectively.

Sensitivity Analyses

Our results were robust across numerous sensitivity analyses that examined alternative outcome measures, added disease and condition variables, altered the composition of the study sample, corrected BMI for self-reporting bias, implemented versions of the standard linear Blinder-Oaxaca decomposition, and adjusted for selection into Medicare Advantage (Appendix Tables S3.4 and S3.5).

DISCUSSION

By using nationally representative survey data from the HRS linked to Medicare claims, we examined whether smoking, alcohol consumption, BMI, and physical activity contribute to geographic variation in Medicare spending and utilization. Our results demonstrated that these factors collectively explained 7% of the difference in spending between higher- and lower-spending regions. While the majority of the difference in spending was not explained by beneficiary characteristics, improving the health behaviors of beneficiaries in higher-spending regions presents an opportunity to modestly reduce geographic variation in spending while also improving the health of the Medicare population.

This study contributes to the ongoing debate over the importance of beneficiaries' characteristics as determinants of geographic variation in Medicare spending. In contrast to prior research, we included a broader set of characteristics, many previously unstudied; these included health behaviors and modifiable risk factors of smoking status, alcohol consumption, BMI, and physical activity, as well as household wealth, poverty status, household size, limitations in ADLs and IADLs, and cognitive status. By focusing on behavioral risk factors, we sought to isolate the importance of beneficiaries' lifestyle characteristics that are potentially modifiable through interventions that promote healthy behaviors in patients or, alternatively, that train and incentivize physicians to help patients quit smoking, consume alcohol within acceptable ranges, maintain healthily diets and weights, and engage in appropriate physical activity.[240] Public health research demonstrates the substantial opportunity to reduce morbidity, mortality, and healthcare costs through such lifestyle changes.[241-244]

The findings of our study are consistent with prior research, which has shown that beneficiary characteristics explain only a moderate amount of geographic variation in Medicare spending. Two prior and related studies, using a different dataset and empirical approach, demonstrated that measures of blood pressure, diabetes, BMI, smoking history, and self-rated health explained 18% of the difference in spending between the highest- and lowest-spending regions,[38] but adding over 10 additional health variables—among these, the presence and diagnosis of specific diseases, and whether beneficiaries died or had proxy respondents—could explain 29% of the regional difference in spending.[37] However, our results differ from two notable and recent studies, which determined that patient characteristics account for the majority of regional variation in healthcare spending. The first implemented a modified version of

Medicare's Hierarchical Condition Categories model—a risk adjustment model based on billing data—and found that population health explains between 75% and 85% of spending variations.[39] The second study used the state as the unit of analysis (rather than the beneficiary), and found that 81% of state-level variation in Medicare spending could be explained by a limited set of aggregate demographic and health measures.[245] The findings of both studies remain susceptible to important biases or methodological limitations,[40-42,44,246] and contradict a diverse literature that emphasizes provider factors as key determinants of geographic variation in spending.[94,96,199] In our study, the inclusion of respondent-reported diseases and conditions, plausibly associated with regional diagnostic practices, did not alter the percentage of variation explained.

When viewed collectively with this prior research, our results reinforce the perspective that non-beneficiary factors are likely to be important determinants of geographic variation in Medicare spending and point toward the role of payment reform and policies aimed at providers' treatment patterns. Nevertheless, behavioral risk factors were an important and distinct contributor amongst the observable characteristics we studied, indicating that these factors should not be ignored when making regional comparisons of spending. While a 7% reduction in Medicare spending in higher-spending regions—such as through smoking cessation programs or interventions to increase physical activity—seems small in an absolute sense, it nonetheless represents approximately $70 million of potential savings based on Medicare expenditures in 2013. Furthermore, policy interventions promoting healthy behaviors within the whole Medicare population may also be important for reducing across-the-board utilization through improvements in health, regardless of the effect these interventions have on geographic differences in spending.[244,247-250] Unlike initiatives targeting provider practice patterns, efforts to improve beneficiaries' health behaviors may also increase life expectancy and enhance quality of life.

Our findings also suggest that alternative policies focused on providers—such as those that would adjust provider reimbursements based on regional spending patterns—could unduly reward or penalize certain regions if beneficiary characteristics are not adequately accounted for. This study highlights the utility of national surveys such as the HRS for investigating the relevance of beneficiary characteristics—characteristics that are rarely available in administrative datasets alone—and for demonstrating the potential for better risk adjustment in the future if health behaviors are collected through electronic health records. Nevertheless, given the relative importance of non-beneficiary characteristics, future research can elucidate how provider factors, such as physicians' training, local treatment norms, and practice patterns, contribute to geographic variation in Medicare spending.

This study has several limitations. First, because our initial study population was selected based on enrollment in traditional (fee-for-service) Medicare, our results may not generalize to Medicare Advantage or privately insured beneficiaries. To reduce bias owing to the exclusion of Medicare Advantage beneficiaries (whose spending and utilization data is not released to researchers), we used 2004 as the base year in our analyses because traditional Medicare accounted for 87% of total Medicare enrollment in 2004—higher than in subsequent years.[224] Sensitivity analyses demonstrated that the observable characteristics of respondents leaving traditional Medicare during the study period were largely similar in higher- versus lower-spending regions (Appendix Table S4.6), increasing the likelihood that unmeasured aspects of health status were also similar.[251] Explicit adjustments for selection into Medicare Advantage did not substantively alter our results.

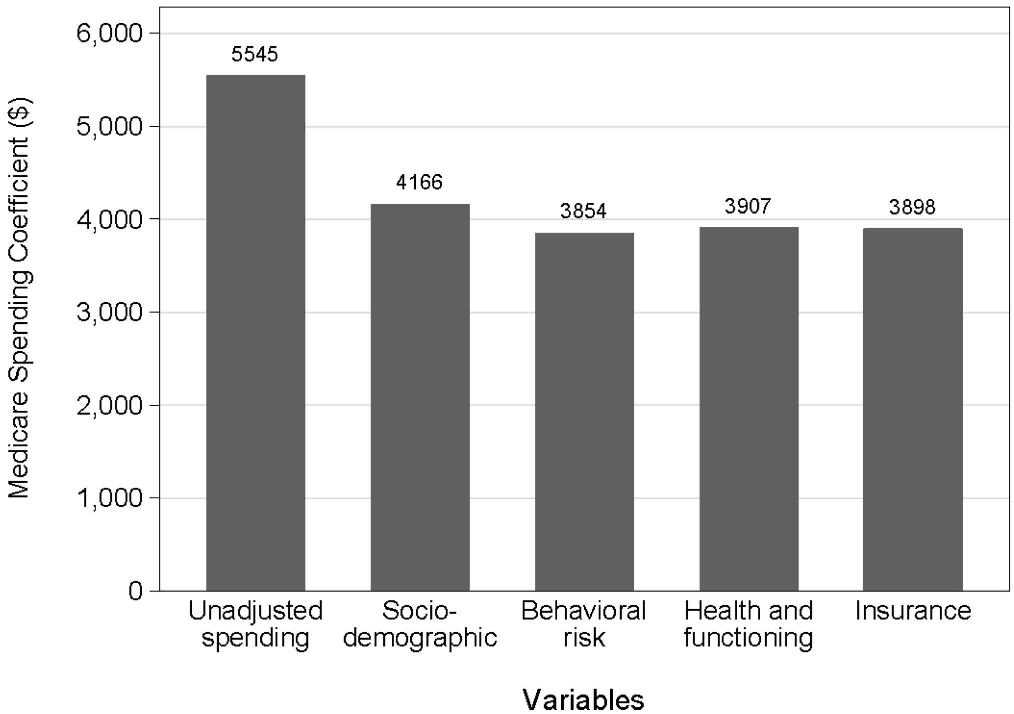
Second, although the difference in spending between higher- and lower-spending regions that was not explained by beneficiary characteristics can be considered analogous to a "treatment effect" of regional practice patterns on beneficiary-level spending,[252] we do not ascribe a causal interpretation to this relationship. Some of the observable variables included in our analyses are unlikely to be pre-treatment variables: for example, a beneficiary's current health and functional status may be affected by healthcare use in prior years.[253] In addition, although we included a comprehensive set of plausible confounders consistent with prior research and theoretical models of health service utilization,[210] the absence of some relevant confounders—particularly those related to health status and severity of illness—from our models may be an extant source of bias. However, the results of sensitivity analyses that included additional disease variables were consistent with our main findings.

Finally, information on behavioral risk factors was obtained by self report and may be subject to measurement error, particularly underreporting of socially undesirable behaviors, misreporting of height and weight, or difficulty recalling the extent of certain behaviors.[88,239,254-256] On balance, we do not expect measurement error to differ in higher- versus lower-spending regions, reducing the possibility that measurement error biased the results of the decomposition analyses. Nevertheless, we implemented a published method to correct BMI for self-reported height and weight,[239] which did not change our results. BMI derived from self-reported height and weight was highly correlated (95%) with BMI based on measured height and weight for a random subsample of respondents for whom HRS collected both sets of measures. HRS measures of behavioral risk factors and their prevalence are also concordant with other national studies such as the National Health Interview Survey and National Health and Nutrition Examination Survey.[257]

Despite these limitations, our study provides evidence of the role of smoking status, alcohol consumption, BMI, and physical activity as determinants of geographic variation in Medicare spending. Comparing Medicare beneficiaries in higher- and lower-spending regions who participated in a large, nationally representative, and longitudinal study, we found that such health behaviors and modifiable risk factors explained 7% of the regional variation in Medicare spending. Health promotion programs targeting older Americans may present an opportunity to modestly reduce geographic variation in healthcare spending, but additional research is needed to identify non-beneficiary characteristics that may be particularly amenable to policy interventions.

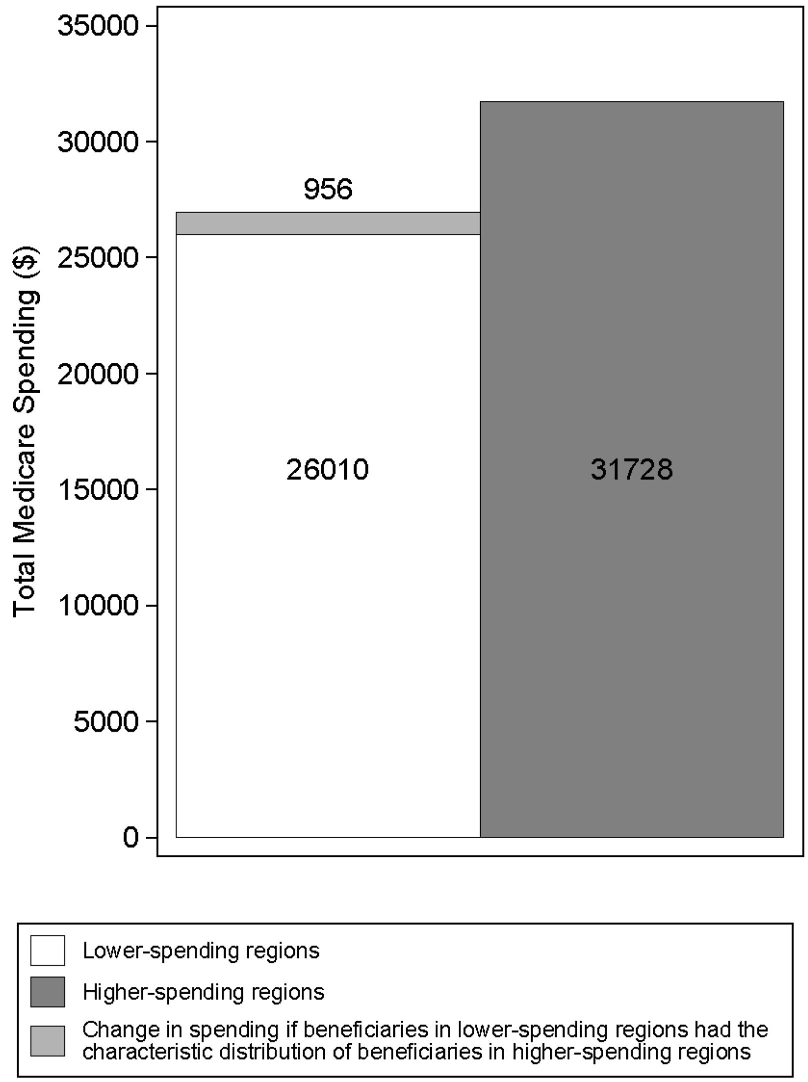
FIGURES

Figure 3.1. Differences in Medicare Spending Between Lower-Spending Regions and Higher-Spending Regions, Based on Five Models



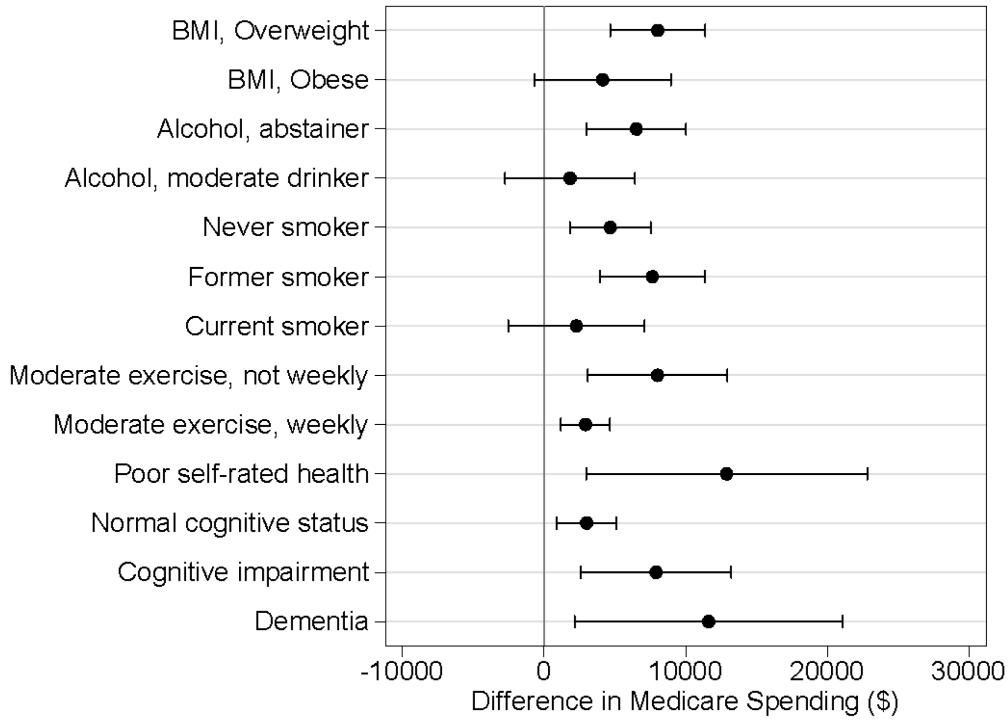
Medicare spending per beneficiary between 2004 and 2006 is price-adjusted and expressed in 2006 U.S. dollars. The coefficient is the average marginal effect from a generalized linear model of a dichotomous dummy variable denoting higher-versus-lower spending hospital referral regions. The first model included only the dummy variable for regional spending levels, representing the difference in price-adjusted Medicare spending between higher- and lower-spending regions. The second model added sociodemographic variables (age, sex, race/ethnicity, educational achievement, household wealth, poverty status, marital status, household size, and urban status); the third model added the behavioral risk factor variables (smoking status, alcohol consumption, body mass index, and physical activity); the fourth model added health and functional status variables (self-rated health status, limitations in activities of daily living and instrumental activities of daily living, cognitive status, proxy respondent status [as an indirect measure of health], and death between 2002-2006); the fifth model added insurance variables (indicators for Medicare Advantage use in 2005 or 2006, discontinuous full Part A and B enrollment between 2004 and 2006, and supplementary private insurance).

Figure 3.2. Total Regression-Adjusted Medicare Spending Comparing Higher- and Lower-Spending Regions



Average predicted Medicare spending is based on average marginal effects from a generalized linear model that adjusts for sociodemographic characteristics, behavioral risk factors, health and functional status characteristics, and insurance coverage. Regional spending level is based on data reported in *The Dartmouth Atlas of Health* Care and refers to total nondrug Part A and B spending in 2004, measured at the hospital referral region (HRR) level. “Lower” refers to HRRs below the median (median, $6,680; range, $5,112 to $7,324) and “higher” refers to HRRs above the median (median, $7,986; range, $7,330 to $11,609).

Figure 3.3. Differences in Regression-Adjusted Medicare Spending Between Higher- and Lower- Spending Regions for Selected Variables



Abbreviations: BMI—body mass index
Predicted Medicare spending is based on average marginal effects from a generalized linear model that adjusts for sociodemographic characteristics, behavioral risk factors, health and functional status characteristics, and insurance coverage. For each variable, the difference (and 95% confidence interval) is calculated as the average predicted spending for beneficiaries in lower-spending regions subtracted from the average predicted spending for beneficiaries in higher-spending regions.

TABLES

Table 3.1. Characteristics of Beneficiaries by Hospital Referral Region (HRR) Spending Level, 2004

		Regional Spending Level[a]			
Variable	**Total**	**Lower**	**Higher**	**Difference**	**P Value**
	(N = 8476)	(N = 4378)	(N = 4098)		
Spending and utilization[b]					
Medicare spending per beneficiary, mean (SD)	27759 (40311)	25078 (36005)	30623 (44279)	5545	<0.001
Inpatient admissions per beneficiary, mean (SD)	1.1 (1.8)	1.0 (1.7)	1.2 (1.9)	0.2	<0.001
Outpatient facility and physician office visits, mean (SD)	32.0 (27.9)	31.4 (27.5)	32.7 (28.2)	1.3	0.027
Sociodemographic characteristics					
Age, mean (SD)	75.2 (7.7)	75.0 (7.6)	75.5 (7.8)	0.4	0.009
Female (%)	57.6	56.7	58.5	1.9	0.08
Race/ethnicity (%)					
White	79.1	81.5	76.5	-5.0	<0.001
Black	12.9	12.2	13.7	1.5	0.04
Hispanic	6.4	4.5	8.3	3.8	<0.001
Other	1.7	1.8	1.5	-0.3	0.25
Educational achievement (%)					
Less than high school	28.6	25.1	32.2	7.1	<0.001
Completed high school	36.9	36.8	37.1	0.3	0.75
More than high school	34.5	38.1	30.7	-7.5	<0.001
Household wealth (%)[c]					
Quintile 1	20.1	18.3	22.0	3.7	<0.001
Quintile 2	19.1	17.3	21.0	3.7	<0.001
Quintile 3	19.7	19.5	19.9	0.4	0.64
Quintile 4	19.8	21.3	18.3	-3.0	0.001
Quintile 5	21.3	23.7	18.8	-4.8	<0.001
Below the federal poverty threshold (%)	10.3	8.7	12.0	3.3	<0.001
Married (%)	55.7	57.1	54.3	-2.8	0.009

Household size, mean (SD)	2.0 (1.0)	1.9 (0.9)	2.0 (1.0)	0.0	0.05
Urban (%)	68.9	66.1	71.8	5.7	<0.001
Behavioral risk factors					
Smoking status (%)					
Nonsmoker	43.5	43.7	43.2	-0.5	0.67
Former smoker	47.1	47.6	46.6	-1.0	0.36
Current smoker	9.5	8.8	10.2	1.5	0.02
Alcohol consumption (%)					
Abstainers	60.6	57.7	63.7	6.0	<0.001
Light drinkers	23.9	26.0	21.8	-4.2	<0.001
Moderate drinkers	11.2	12.0	10.4	-1.6	0.02
Heavy drinkers[d]	4.2	4.4	4.1	-0.2	0.62
Body mass index (BMI) category (%)					
Underweight	3.0	2.8	3.3	0.4	0.27
Normal weight	37.6	37.5	37.8	0.4	0.73
Overweight	37.2	37.3	37.0	-0.3	0.77
Obese	22.1	22.3	21.9	-0.5	0.60
Physical activity (weekly)					
Light exercise (%)	74.0	75.7	72.2	-3.4	<0.001
Moderate exercise (%)	60.9	62.7	58.9	-3.8	<0.001
Vigorous exercise (%)	24.5	25.0	23.9	-1.1	0.24
Health and functional status					
Self-rated health (%)					
Excellent	8.1	8.9	7.3	-1.6	0.006
Very good	25.1	27.7	22.2	-5.5	<0.001
Good	32.3	31.7	32.9	1.2	0.241
Fair	23.3	21.7	24.9	3.2	<0.001
Poor	11.3	10.0	12.7	2.8	<0.001
ADL limitations, mean (SD)	0.5 (1.2)	0.5 (1.1)	0.5 (1.2)	0.1	0.01
IADL limitations, mean (SD)	0.5 (1.2)	0.5 (1.2)	0.5 (1.2)	0.1	0.04

Cognitive functioning (%)					
Normal cognition	67.0	68.9	65.1	-3.8	<0.001
Cognitive impairment without dementia	21.7	20.6	22.9	2.3	0.009
Dementia	11.3	10.6	12.1	1.5	0.03
Proxy respondent (%)	10.4	9.5	11.5	2.0	0.002
Death in 2004-2006 (%)	12.9	12.9	13.0	0.1	0.92
Insurance coverage					
Medicare Advantage use in 2005 or 2006 (%)	8.1	8.0	8.2	0.2	0.76
Discontinuous full Part A and B enrollment, 2004-2006 (%)	18.1	18.2	17.9	-0.3	0.73
Supplementary private insurance (%)	68.3	70.8	65.7	-5.1	<0.001

Abbreviations: ADL, activity of daily living; IADL, instrumental activity of daily living.

[a] Regional spending level is based on data reported in *The Dartmouth Atlas of Health* Care and refers to total nondrug Part A and B spending in 2004, measured at the hospital referral region (HRR) level. Beneficiaries were divided into two mutually exclusive categories of higher- versus lower-regional spending on the basis of the spending in their HRR in 2004: HRRs below the median designate lower-spending regions (median, $6,680; range, $5,112 to $7,324) and HRRs above the median designate higher-spending regions (median, $7,986; range, $7,330 to $11,609). Percentages may not add up exactly to 100 due to rounding.

[b] Spending and utilization outcomes are cumulative over the 2004-2006 period. Medicare spending includes total nondrug Part A and B reimbursements as well as the amount contributed by beneficiaries; spending was adjusted for cross-sectional differences in prices and is expressed in 2006 U.S. dollars based on the medical care component of the Consumer Price Index.

[c] Respondents' total household wealth (excluding individual retirement accounts) was measured in quintiles based on the full wealth distribution of respondents in the corresponding HRS survey wave.

[d] Heavy and very heavy drinkers were collapsed into 1 category to protect the confidentiality of beneficiaries.

Table 3.2. Decomposition of Differences in Medicare Spending and Utilization by Regional Spending Level[a]

	Total Spending	Inpatient Admissions	Outpatient Facility and Physician Office Visits
Difference between higher- and lower-spending regions[b]	$5,718	0.19	1.31
Percentage explained			
Difference attributable to beneficiary characteristics, %	17	24	1
Difference not attributable to beneficiary characteristics, %	83	76	99
Difference attributable to, %			
Sociodemographic characteristics[c]	-1	-1	-22
Behavioral risk factors[d]	7	11	-6
Health status[e]	14	18	56
Functional status[f]	3	1	-4
Cognitive status	-1	1	-8
Insurance[g]	-6	-6	-16

[a] Spending and utilization outcomes are cumulative over the 2004-2006 period. Medicare spending includes total nondrug Part A and B reimbursements as well as the amount contributed by beneficiaries. Positive percentages indicate that the variable set contributes toward increasing the difference in spending or utilization between higher- and lower-spending regions whereas negative percentages indicate that the variable set decreases the difference. Percentages may not add up exactly to 100 due to rounding.

[b] Differences are based on average predicted outcomes from generalized linear models used in the decomposition. Regional spending level is based on data reported in *The Dartmouth Atlas of Health* Care and refers to total nondrug Part A and B spending in 2004, measured at the hospital referral region (HRR) level. Beneficiaries were divided into two mutually exclusive categories of higher- versus lower-regional spending on the basis of the spending in their HRR in 2004: HRRs below the median designate lower-spending regions (median, $6,680; range, $5,112 to $7,324) and HRRs above the median designate higher-spending regions (median, $7,986; range, $7,330 to $11,609).

[c] Sociodemographic characteristics include age, sex, race/ethnicity, educational achievement, household wealth, poverty status, marital status, household size, and urban status.

[d] Behavioral risk factors include smoking status, alcohol consumption, body mass index, and physical activity.

[e] Health status includes self-rated health, proxy respondent status (as an indirect measure of health), and death between 2004-2006.

[f] Functional status includes limitations in activities of daily living and instrumental activities of daily living.

[g] Insurance includes indicators for Medicare Advantage use in 2005 or 2006, discontinuous full Part A and B enrollment between 2004 and 2006, and supplementary private insurance.

APPENDIX

Table S3.1. Predicted Spending for Beneficiaries' Behavioral Risk Factors and Health Associated with Regional Spending Levels

	Predicted Spending, $ (95% CI)[a]			
Variable	**Lower Spending Regions**	**Higher Spending Regions**	**Difference[b]**	**P Value**
Smoking status				
Never smoker	23899 (22201 to 25597)	28565 (26267 to 30863)	4666 (1809 to 7523)	0.00
Former smoker	28581 (26591 to 30572)	36218 (33099 to 39337)	7637 (3937 to 11337)	0.00
Current smoker	22460 (19420 to 25500)	24740 (21016 to 28464)	2280 (-2527 to 7088)	0.35
Alcohol consumption				
Abstainer	30069 (27896 to 32243)	36573 (33812 to 39334)	6504 (2989 to 10018)	0.00
Light drinker	21260 (19142 to 23378)	23986 (21821 to 26150)	2725 (-303 to 5754)	0.08
Moderate drinker	20652 (17061 to 24242)	22485 (19623 to 25346)	1833 (-2758 to 6424)	0.43
Heavy drinker	16788 (12935 to 20640)	21646 (15181 to 28111)	4858 (-2668 to 12384)	0.21
Very heavy drinker	12984 (2020 to 23949)	27617 (6754 to 48480)	14633 (-8935 to 38201)	0.22
Body mass index				
Underweight	38726 (24857 to 52595)	44529 (35261 to 53798)	5803 (-10878 to 22485)	0.50
Normal	27393 (24869 to 29917)	31613 (28961 to 34265)	4220 (558 to 7881)	0.02
Overweight	22238 (20636 to 23840)	30262 (27321 to 33202)	8024 (4675 to 11372)	0.00
Obese	28414 (25503 to 31324)	32541 (28714 to 36369)	4128 (-681 to 8936)	0.09
Light exercise				
Less than weekly	44991 (41242 to 48741)	53240 (47995 to 58486)	8249 (1801 to 14697)	0.01
At least weekly	20086 (19024 to 21147)	23721 (22094 to 25347)	3635 (1693 to 5577)	0.00
Moderate exercise				
Less than weekly	38455 (35565 to 41345)	46454 (42466 to 50442)	7999 (3073 to 12924)	0.00
At least weekly	18777 (17721 to 19833)	21694 (20331 to 23056)	2917 (1193 to 4641)	0.00
Vigorous exercise				
Less than weekly	29282 (27466 to 31098)	35727 (33240 to 38213)	6445 (3366 to 9524)	0.00
At least weekly	16340 (14685 to 17995)	19232 (17416 to 21049)	2892 (435 to 5350)	0.02
Self-rated health status				

Excellent health	12177 (10209 to 14145)	13591 (10812 to 16369)	1414 (-1991 to 4819)	0.42
Very good health	14886 (13475 to 16296)	17688 (15833 to 19543)	2802 (472 to 5133)	0.02
Good health	22328 (20391 to 24264)	25513 (23755 to 27271)	3185 (570 to 5800)	0.02
Fair health	39094 (36016 to 42171)	40703 (36158 to 45247)	1609 (-3879 to 7097)	0.57
Poor health	53668 (47554 to 59783)	66549 (58733 to 74365)	12881 (2957 to 22804)	0.01
Cognitive status				
Normal cognition	21689 (20402 to 22976)	24683 (23018 to 26348)	2994 (890 to 5098)	0.01
Cognitive impairment	31511 (28656 to 34367)	39413 (34943 to 43883)	7902 (2598 to 13206)	0.00
Dementia	45122 (38601 to 51643)	56737 (49883 to 63591)	11616 (2155 to 21076)	0.02
No proxy	24360 (23046 to 25674)	28979 (26953 to 31005)	4619 (2204 to 7034)	0.00
Proxy respondent	42893 (36076 to 49709)	54049 (47764 to 60335)	11156 (1884 to 20429)	0.02
Survived	20150 (19113 to 21187)	24856 (23029 to 26682)	4706 (2606 to 6806)	0.00
Death 2004-2006	66781 (59662 to 73901)	79522 (71048 to 87995)	12740 (1673 to 23808)	0.02

[a] Predicted Medicare spending is based on average marginal effects from a generalized linear model that adjusts for sociodemographic characteristics, behavioral risk factors, health and functional status characteristics, and insurance coverage.
[b] For each variable, the difference (and 95% confidence interval) is calculated as the average predicted spending for beneficiaries in lower-spending regions subtracted from the average predicted spending for beneficiaries in higher-spending regions.

Table S3.2. Regression Models of Medicare Spending and Utilization for Beneficiaries in Lower Spending Regions[a]

	Total Spending, Coefficient (SE)	Inpatient Admissions, Coefficient (SE)	Outpatient Facility and Physician Office Visits, Coefficient (SE)
Age 70 to 74	0.15 (0.05) **	0.23 (0.08) **	0.027 (0.04)
Age 75 to 79	0.27 (0.06) ***	0.37 (0.10) ***	0.16 (0.04) ***
Age 80 to 84	0.22 (0.08) **	0.37 (0.09) ***	0.16 (0.05) **
Age 85 to 89	0.22 (0.07) **	0.39 (0.09) ***	0.15 (0.05) **
Age 90 and over	0.032 (0.1)	0.23 (0.1)	-0.092 (0.07)
Female	-0.034 (0.06)	-0.040 (0.06)	0.16 (0.03) ***
African American	-0.033 (0.08)	0.017 (0.09)	-0.089 (0.05)
Hispanic	-0.15 (0.1)	-0.28 (0.1) *	-0.13 (0.09)
Other race	-0.11 (0.2)	-0.17 (0.2)	-0.092 (0.1)
Less than high school	0.048 (0.07)	0.11 (0.08)	-0.0043 (0.04)
More than high school	0.093 (0.06)	0.14 (0.06) *	0.074 (0.04)
Wealth quintile 1	0.12 (0.06)	0.058 (0.08)	0.10 (0.05)
Wealth quintile 2	-0.065 (0.06)	-0.089 (0.09)	-0.041 (0.04)
Wealth quintile 4	-0.078 (0.07)	-0.12 (0.08)	0.0097 (0.05)
Wealth quintile 5	-0.028 (0.06)	-0.14 (0.07) *	0.039 (0.05)
Below poverty threshold	-0.12 (0.08)	-0.18 (0.08) *	0.092 (0.06)
Married	-0.10 (0.06)	-0.098 (0.06)	0.058 (0.04)
Household size	0.012 (0.03)	0.011 (0.03)	-0.0026 (0.02)
Urban	0.0051 (0.05)	-0.060 (0.06)	-0.058 (0.06)
BMI Underweight	-0.017 (0.2)	-0.15 (0.2)	-0.24 (0.09) **
BMI Overweight	-0.052 (0.05)	-0.079 (0.05)	0.052 (0.03)
BMI Obese	0.11 (0.05) *	0.057 (0.06)	0.080 (0.03) *
Light drinker	-0.080 (0.06)	-0.17 (0.06) **	0.023 (0.03)
Moderate drinker	-0.077 (0.10)	-0.13 (0.10)	0.016 (0.05)
Heavy drinker	-0.21 (0.1)	-0.24 (0.2)	-0.085 (0.08)

Very heavy drinker	-0.57 (0.4)	-1.28 (0.6) *	-0.73 (0.2) **
Former smoker	0.12 (0.04) **	0.057 (0.05)	0.049 (0.03)
Current smoker	-0.022 (0.07)	-0.064 (0.09)	-0.25 (0.06) ***
Weekly vigorous exercise	-0.053 (0.07)	-0.16 (0.08) *	-0.0018 (0.04)
Weekly moderate exercise	-0.14 (0.05) **	-0.13 (0.05) *	-0.059 (0.03)
Weekly light exercise	-0.24 (0.06) ***	-0.23 (0.07) ***	-0.057 (0.04)
Excellent health	-0.87 (0.1) ***	-0.97 (0.1) ***	-0.67 (0.07) ***
Very good health	-0.72 (0.07) ***	-0.85 (0.10) ***	-0.51 (0.05) ***
Good health	-0.43 (0.07) ***	-0.49 (0.08) ***	-0.34 (0.05) ***
Fair health	-0.11 (0.06)	-0.17 (0.08) *	-0.10 (0.05) *
ADL limitations	0.049 (0.03)	0.050 (0.03)	-0.026 (0.02)
IADL limitations	0.055 (0.03)	-0.022 (0.03)	0.0035 (0.02)
Cognitive impairment without dementia	-0.0034 (0.05)	0.12 (0.05) *	-0.044 (0.03)
Dementia	-0.18 (0.08) *	-0.056 (0.10)	-0.15 (0.06) *
Proxy respondent	-0.15 (0.09)	-0.18 (0.1)	-0.24 (0.07) ***
Death 2004-06	1.27 (0.1) ***	1.16 (0.1) ***	0.40 (0.10) ***
Any Medicare Advantage use in 2005-06	-0.41 (0.09) ***	-0.39 (0.1) ***	-0.34 (0.06) ***
Discontinuous full Part A and B enrollment, 2004-06	-0.71 (0.1) ***	-0.47 (0.1) ***	-0.72 (0.09) ***
Supplemental private insurance	0.22 (0.05) ***	0.21 (0.07) **	0.14 (0.04) ***

Abbreviations: BMI, body mass index; ADL, activity of daily living; IADL, instrumental activity of daily living.

[a] The model for spending was estimated using a generalized linear model (GLM) with a log link and gamma distribution; the models for utilization outcomes used a negative binomial distribution. Spending and utilization outcomes are cumulative over the 2004-2006 period. Medicare spending includes total nondrug Part A and B reimbursements as well as the amount contributed by beneficiaries. Standard errors (SEs) were clustered on hospital referral regions.

* for p<.05; ** for p<.01; and *** for p<.001. N = 4233.

Table S3.3. Regression Models of Medicare Spending and Utilization for Beneficiaries in Higher Spending Regions[a]

	Total Spending, Coefficient (SE)	Inpatient Admissions, Coefficient (SE)	Outpatient Facility and Physician Office Visits, Coefficient (SE)
Age 70 to 74	0.090 (0.05)	0.077 (0.07)	0.076 (0.04)
Age 75 to 79	0.24 (0.06) ***	0.28 (0.07) ***	0.098 (0.04) *
Age 80 to 84	0.27 (0.07) ***	0.30 (0.08) ***	0.12 (0.05) *
Age 85 to 89	0.16 (0.08)	0.27 (0.09) **	-0.021 (0.05)
Age 90 and over	0.068 (0.08)	0.20 (0.10) *	-0.11 (0.06)
Female	-0.12 (0.05) *	-0.14 (0.07) *	0.14 (0.03) ***
African American	0.057 (0.07)	-0.038 (0.08)	-0.019 (0.05)
Hispanic	0.12 (0.1)	-0.23 (0.07) **	0.11 (0.06)
Other race	0.0027 (0.2)	0.027 (0.2)	-0.23 (0.1)
Less than high school	0.0084 (0.05)	0.014 (0.06)	-0.023 (0.04)
More than high school	-0.0030 (0.05)	-0.079 (0.06)	0.00019 (0.03)
Wealth quintile 1	-0.042 (0.07)	-0.10 (0.08)	0.020 (0.05)
Wealth quintile 2	-0.065 (0.06)	-0.099 (0.06)	-0.087 (0.04) *
Wealth quintile 4	-0.057 (0.06)	-0.15 (0.09)	-0.011 (0.04)
Wealth quintile 5	0.077 (0.06)	-0.059 (0.08)	0.097 (0.04) *
Below poverty threshold	0.15 (0.06) *	0.026 (0.08)	0.099 (0.05)
Married	-0.078 (0.06)	-0.14 (0.07) *	0.085 (0.04) *
Household size	-0.0087 (0.02)	0.036 (0.02)	-0.030 (0.01) *
Urban	0.13 (0.06) *	0.016 (0.07)	0.031 (0.04)
BMI Underweight	-0.12 (0.1)	-0.015 (0.2)	-0.22 (0.08) **
BMI Overweight	0.064 (0.06)	0.090 (0.06)	0.026 (0.03)
BMI Obese	0.10 (0.07)	0.090 (0.07)	0.097 (0.04) *
Light drinker	-0.13 (0.05) **	-0.13 (0.05) **	-0.028 (0.03)
Moderate drinker	-0.099 (0.07)	-0.23 (0.08) **	-0.030 (0.05)
Heavy drinker	-0.13 (0.1)	-0.22 (0.1)	-0.13 (0.08)

Very heavy drinker	-0.12 (0.3)	-0.0074 (0.4)	-0.033 (0.2)
Former smoker	0.18 (0.04) ***	0.22 (0.05) ***	0.068 (0.03) *
Current smoker	-0.097 (0.08)	0.011 (0.09)	-0.20 (0.04) ***
Weekly vigorous exercise	-0.076 (0.06)	-0.14 (0.07)	-0.039 (0.03)
Weekly moderate exercise	-0.20 (0.05) ***	-0.18 (0.06) **	-0.086 (0.03) *
Weekly light exercise	-0.16 (0.06) *	-0.19 (0.06) **	-0.052 (0.04)
Excellent health	-0.84 (0.1) ***	-0.87 (0.2) ***	-0.78 (0.1) ***
Very good health	-0.66 (0.08) ***	-0.81 (0.08) ***	-0.49 (0.07) ***
Good health	-0.41 (0.07) ***	-0.50 (0.07) ***	-0.29 (0.06) ***
Fair health	-0.15 (0.07) *	-0.30 (0.07) ***	-0.096 (0.05)
ADL limitations	0.098 (0.02) ***	0.023 (0.02)	0.013 (0.02)
IADL limitations	0.024 (0.03)	0.069 (0.03) *	-0.051 (0.02) *
Cognitive impairment without dementia	0.095 (0.05)	0.15 (0.07) *	-0.053 (0.03)
Dementia	-0.036 (0.09)	-0.066 (0.10)	-0.11 (0.07)
Proxy respondent	-0.19 (0.07) **	-0.20 (0.08) *	-0.15 (0.04) ***
Death 2004-06	1.29 (0.1) ***	0.99 (0.1) ***	0.52 (0.1) ***
Any Medicare Advantage use in 2005-06	-0.30 (0.09) **	-0.14 (0.09)	-0.27 (0.06) ***
Discontinuous full Part A and B enrollment, 2004-06	-0.75 (0.1) ***	-0.45 (0.1) ***	-0.90 (0.1) ***
Supplemental private insurance	0.093 (0.05)	0.045 (0.06)	0.15 (0.04) ***

Abbreviations: BMI, body mass index; ADL, activity of daily living; IADL, instrumental activity of daily living.

[a] The model for spending was estimated using a generalized linear model (GLM) with a log link and gamma distribution; the models for utilization outcomes used a negative binomial distribution. Spending and utilization outcomes are cumulative over the 2004-2006 period. Medicare spending includes total nondrug Part A and B reimbursements as well as the amount contributed by beneficiaries. Standard errors (SEs) were clustered on hospital referral regions.

* for $p<.05$; ** for $p<.01$; and *** for $p<.001$. N = 3985.

Sensitivity Analyses

The following sensitivity analyses are presented in Appendix Tables S3.4 and S3.5. Sensitivity analyses A through H implemented the decomposition based on generalized linear models, as used in the main analysis. To assess the robustness of our results, we:

A. Examined 5-year Medicare spending beginning in 2004 (rather than 3 as in main analysis)
B. Used 2006 as a base year (instead of 2004), and examined 3-year Medicare spending through 2008
C. Used price-unadjusted measure of Medicare spending
D. Included disease/condition variables generally believed to be associated with regional variation in diagnostic practices as additional set of independent variables
E. Excluded veterans
F. Excluded respondents who did not have continuous Part A and B enrollment over the observation period
G. Excluded respondents who died after the base year
H. Corrected body mass index for respondent self-report using the Cawley method described in the main text
I. Implemented the standard Blinder-Oaxaca decomposition (Medicare spending measured in dollars)
J. Implemented the standard Blinder-Oaxaca decomposition (Medicare spending measured in log dollars)
K. Adjusted for Medicare Advantage selection (implemented with standard Blinder-Oaxaca decomposition)

Table S3.4. Decomposition of Differences in Medicare Spending by Regional Spending Level, Sensitivity Analyses A through F[a]

	Main	A	B	C	D	E	F
Difference between higher- and lower-spending regions, $[b]	5,718	11,556	8,076	5,820	5,664	5,370	5,577
Percentage explained							
Difference attributable to beneficiary characteristics, %	17	12	16	18	16	17	20
Difference not attributable to beneficiary characteristics, %	83	88	84	82	84	83	80
Difference attributable to, %							
Sociodemographic characteristics[c]	-1	-1	-4	1	-1	1	-3
Behavioral risk factors[d]	7	4	5	7	7	8	8
Health status[e]	14	14	17	14	10	12	16
Functional status[f]	3	1	2	3	3	3	3
Cognitive status	-1	-1	1	-1	-1	-2	0
Insurance[g]	-6	-5	-5	-6	-6	-4	-4
Diseases/conditions[h]	--	--	--	--	5	--	--

[a] Medicare spending includes total nondrug Part A and B reimbursements as well as the amount contributed by beneficiaries. Positive percentages indicate that the variable set contributes toward increasing the difference in spending between higher- and lower-spending regions whereas negative percentages indicate that the variable set decreases the difference. Percentages may not add up exactly to 100 due to rounding.

[b] Differences are based on average predicted outcomes from generalized linear models used in the decomposition. Regional spending level is based on data reported in *The Dartmouth Atlas of Health* Care and refers to total nondrug Part A and B spending in 2004, measured at the hospital referral region (HRR) level. Beneficiaries were divided into two mutually exclusive categories of higher- versus lower-regional spending on the basis of the spending in their HRR in 2004: HRRs below the median designate lower-spending regions (median, $6,680; range, $5,112 to $7,324) and HRRs above the median designate higher-spending regions (median, $7,986; range, $7,330 to $11,609).

[c] Sociodemographic characteristics include age, sex, race/ethnicity, educational achievement, household wealth, poverty status, marital status, household size, and urban status.

[d] Behavioral risk factors include smoking status, alcohol consumption, body mass index, and physical activity.

[e] Health status includes self-rated health, proxy respondent status (as an indirect measure of health), and death between 2004-2006.

[f] Functional status includes limitations in activities of daily living and instrumental activities of daily living.

[g] Insurance includes indicators for Medicare Advantage use in 2005 or 2006, discontinuous full Part A and B enrollment between 2004 and 2006, and supplementary private insurance.

[h] Diseases/conditions includes a variable that sums the number of diseases as well as indicators for high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, psychiatric condition, and arthritis.

Table S3.5. Decomposition of Differences in Medicare Spending by Regional Spending Level, Sensitivity Analyses G through K[a]

	Main	G	H	I	J	K
Difference between higher- and lower-spending regions, $[b]	5,718	5,077	5,681	5,528	0.25	8,511
Percentage explained						
Difference attributable to beneficiary characteristics, %	17	25	18	17	13	16
Difference not attributable to beneficiary characteristics, %	83	75	82	83	87	84
Difference attributable to, %						
Sociodemographic characteristics[c]	-1	-2	-1	-3	-6	3
Behavioral risk factors[d]	7	9	8	7	4	4
Health status[e]	14	19	15	16	22	9
Functional status[f]	3	4	4	4	2	3
Cognitive status	-1	-1	-1	-2	1	-1
Insurance[g]	-6	-4	-6	-5	-10	-2

[a] Medicare spending includes total nondrug Part A and B reimbursements as well as the amount contributed by beneficiaries. Sensitivity Analysis J reports differences in log dollars (based on natural logarithm). The difference of $5,528 in Sensitivity Analysis I does not equal the unadjusted difference in Medicare spending of $5,545 as expected in the standard Blinder-Oaxaca decomposition because the regressions rely on a complete case analysis and a limited number of observations are dropped due to missing data. The selection model used in Sensitivity Analysis K models beneficiary selection into a Medicare Advantage plan in 2005 or 2006 as a function of the remaining beneficiary characteristics used in the decomposition regressions, in addition to state-level Medicare Advantage access rates as an additional variable (from the Kaiser Family Foundation) to satisfy the exclusion restriction. Positive percentages indicate that the variable set contributes toward increasing the difference in spending between higher- and lower-spending regions whereas negative percentages indicate that the variable set decreases the difference. Percentages may not add up exactly to 100 due to rounding.
[b] Differences are based on average predicted outcomes from generalized linear models used in the decomposition with the exception of Sensitivity Analyses I, J, and K which were based on the standard Blinder-Oaxaca decomposition. Regional spending level is based on data reported in *The Dartmouth Atlas of Health* Care and refers to total nondrug Part A and B spending in 2004, measured at the hospital referral region (HRR) level. Beneficiaries were divided into two mutually exclusive categories of higher- versus lower-regional spending on the basis of the spending in their HRR in 2004: HRRs below the median designate lower-spending regions (median, $6,680; range, $5,112 to $7,324) and HRRs above the median designate higher-spending regions (median, $7,986; range, $7,330 to $11,609).
[c] Sociodemographic characteristics include age, sex, race/ethnicity, educational achievement, household wealth, poverty status, marital status, household size, and urban status.
[d] Behavioral risk factors include smoking status, alcohol consumption, body mass index, and physical activity.
[e] Health status includes self-rated health, proxy respondent status (as an indirect measure of health), and death between 2004-2006.
[f] Functional status includes limitations in activities of daily living and instrumental activities of daily living.
[g] Insurance includes indicators for Medicare Advantage use in 2005 or 2006, discontinuous full Part A and B enrollment between 2004 and 2006, and supplementary private insurance.

Table S3.6. Comparison of Medicare Advantage Beneficiaries in Higher- versus Lower-Spending Regions

		Regional Spending Level[a]			
Variable	**Total**	**Lower**	**Higher**	**Difference**	**P Value**
	(N = 687)	(N = 351)	(N = 336)		
Age, mean (SD)	73.6 (7.1)	74.0 (7.2)	73.3 (6.9)	-0.7	0.230
Female (%)	55.2	53.6	56.8	3.3	0.388
Race/ethnicity (%)					
White	64.3	71.2	57.1	-14.1	0.000
Black	22.9	19.1	26.8	7.7	0.016
Hispanic	12.1	9.4	14.9	5.5	0.028
Other	0.7	0.3	1.2	0.9	0.163
Educational achievement (%)					
Less than high school	39.6	34.5	44.9	10.5	0.005
Complete high school	34.6	37.0	32.1	-4.9	0.178
More than high school	25.8	28.5	22.9	-5.6	0.095
Household wealth (%)[b]					
Quintile 1	27.1	25.1	29.2	4.1	0.228
Quintile 2	24.2	23.1	25.3	2.2	0.497
Quintile 3	19.8	18.8	20.8	2.0	0.505
Quintile 4	15.7	18.5	12.8	-5.7	0.040
Quintile 5	13.2	14.5	11.9	-2.6	0.311
Below poverty threshold (%)	15.7	13.1	18.5	5.3	0.054
Married (%)	59.4	62.1	56.5	-5.6	0.138
Household size, mean (SD)	2.1 (1.2)	2.1 (1.2)	2.2 (1.3)	0.1	0.455
Urban (%)	80.1	76.1	84.2	8.2	0.007
Body mass index (BMI) category (%)					
Underweight	2.4	2.3	2.4	0.1	0.934
Normal	31.8	33.1	30.3	-2.8	0.432
Overweight	39.7	39.8	39.6	-0.1	0.972
Obese	26.2	24.8	27.6	2.8	0.400
Alcohol consumption (%)					
Abstainers	62.1	59.6	64.8	5.2	0.163

Light drinkers	23.7	26.1	21.2	-4.9	0.134
Moderate drinkers	8.9	9.7	8.1	-1.7	0.441
Heavy drinkers[c]	5.3	4.6	6.0	1.4	0.418
Smoking status (%)					
Nonsmoker	43.9	45.7	42.1	-3.6	0.340
Former smoker	42.2	42.3	42.1	-0.2	0.959
Current smoker	13.9	12.0	15.8	3.8	0.149
Weekly vigorous exercise (%)	22.7	24.6	20.8	-3.7	0.244
Weekly moderate exercise (%)	62.1	66.6	57.4	-9.1	0.014
Weekly light exercise (%)	76.4	78.3	74.4	-3.9	0.224
Self-rated health (%)					
Excellent	5.8	6.8	4.8	-2.1	0.246
Very good	22.9	26.8	18.8	-8.0	0.012
Good	32.6	33.0	32.1	-0.9	0.801
Fair	25.5	22.2	28.9	6.6	0.046
Poor	13.2	11.1	15.5	4.4	0.092
ADL limitations, mean (SD)	0.4 (1.0)	0.4 (1.0)	0.4 (1.1)	0.1	0.295
IADL limitations, mean (SD)	0.4 (1.0)	0.4 (1.0)	0.4 (0.9)	0.0	0.681
Cognitive functioning (%)					
Normal	60.2	61.1	59.2	-1.9	0.609
Cognitive impairment without dementia	28.4	28.6	28.3	-0.3	0.931
Dementia	11.4	10.3	12.5	2.2	0.362
Discontinuous full Part A and B enrollment, 2004-06 (%)	9.3	10.3	8.3	-1.9	0.387
Supplementary private insurance (%)	52.8	56.2	49.3	-6.9	0.070

Abbreviations: ADL, activity of daily living; IADL, instrumental activity of daily living.

[a] Regional spending level is based on data reported in *The Dartmouth Atlas of Health* Care and refers to total nondrug Part A and B spending in 2004, measured at the hospital referral region (HRR) level. Beneficiaries were divided into two mutually exclusive categories of higher- versus lower-regional spending on the basis of the spending in their HRR in 2004: HRRs below the median designate lower-spending regions (median, $6,680; range, $5,112 to $7,324) and HRRs above the median designate higher-spending regions (median, $7,986; range, $7,330 to $11,609). Percentages may not add up exactly to 100 due to rounding.

[b] Respondents' total household wealth (excluding individual retirement accounts) was measured in quintiles based on the full wealth distribution of respondents in the corresponding HRS survey wave.

[c] Heavy and very heavy drinkers were collapsed into 1 category to protect the confidentiality of beneficiaries.

CHAPTER FOUR (MANUSCRIPT #3)

ASSOCIATION BETWEEN MEDICARE SPENDING AND BENEFICIARIES' HEALTH, FUNCTIONING, AND MORTALITY

by

Kurt Richard Herzer

ABSTRACT

Importance: There is mixed evidence that higher Medicare spending benefits patients. Little is known about whether the intensity of healthcare spending during and after hospitalization impacts patients' health or functioning.

Objective: To determine whether differences in healthcare spending are associated with beneficiaries' health, functioning, and mortality after hospitalization.

Design: Survey data from the nationally representative Health and Retirement Study were linked to Medicare claims and hospital referral region spending characteristics from the *Dartmouth Atlas of Health Care*. Instrumental variables regression was used to adjust for observed and unobserved differences in beneficiaries' characteristics and to assess the effects of Medicare spending during the year following hospitalization on beneficiaries' health, functioning, and mortality.

Setting: Care provided in US hospital referral regions.

Participants: Acutely ill fee-for-service Medicare beneficiaries hospitalized between 2003 and 2010 for acute myocardial infarction, hip fracture, gastrointestinal bleeding, stroke, cancer, congestive heart failure, pneumonia, acute respiratory failure, and unstable angina.

Exposures: Price-adjusted Medicare Part A and B spending during the year following hospitalization.

Main Outcomes and Measures: Self-rated health status, limitations in activities of daily living (ADLs) and instrumental ADLs (IADLs), pain, cognitive functioning, depressive symptoms, and 1-year mortality. Outcomes were assessed through 2013.

Results: 2,772 Medicare beneficiaries experienced 4,685 hospitalizations. After adjusting for confounding due to health status, a 10% increase in price-adjusted, 1-year Medicare spending was associated with reductions in the probability of new IADL limitations (-1.96 percentage points; 95% confidence interval [CI], -3.88 to -0.03; P=0.05), new depressive symptoms (-2.31

percentage points; 95% CI, -4.04 to -0.59; P=0.009), and mortality (-2.02 percentage points; 95% CI, -3.57 to -0.46; P=0.01). There was no association between higher Medicare spending and self-rated health status, ADL limitations, pain, or cognitive functioning.

Conclusions: Higher Medicare spending in the year following hospitalization was associated with fewer new functional limitations, depressive symptoms, and lower mortality rates, but was not related to other measures of health. While prior research has focused on mortality, we find important relationships between Medicare spending intensity and previously unstudied domains of beneficiaries' health.

INTRODUCTION

Marked variation in Medicare spending across regions of the United States has prompted intense policy debate as to whether the intensity of Medicare spending is related to beneficiaries' health outcomes.[123] Early studies found that Medicare patients in higher-spending regions and hospitals did not have lower mortality rates and experienced few differences in other measures of health system performance.[4,5,159] More recent studies have utilized natural experiments to address the concern that unmeasured differences in patients' health may bias estimates of the relationship between spending and outcomes. These studies found that higher acute care inpatient spending was associated with improved survival.[127,152,258] However, this growing literature lacks a comprehensive assessment of outcomes other than mortality that are important to patients, providers, and policymakers, such as measures of health and functioning.

Motivated by a comprehensive report from the Institute of Medicine highlighting post-acute care as a major driver of variation in spending[3] and the paucity of evidence relating spending to functional status, we examined the relationship between Medicare spending in the year following hospitalization for acute illness and beneficiaries' self-rated health status, functional limitations, pain, cognitive status, and depressive symptoms.

METHODS

Study Data, Population, and Period

We used data from the Health and Retirement Study (HRS), a large, nationally representative, prospectively collected cohort study of older Americans that has been used to study geographic variation in Medicare spending.[199,222,223,259,260] We linked survey measures of socioeconomic status, and health and functioning before and after hospitalization, to Medicare claims. During our study period, 91% of eligible HRS respondents consented to release their Medicare claims for research purposes.

Because our research design required exact Medicare spending, we restricted our study population to respondents who were 65 years or older (N=8,208) and continuously enrolled in fee-for-service Medicare during the 365-day period starting with the index admission (N=7,121). To minimize confounding due to unobserved differences in patient health status, we studied a cohort of acutely ill respondents who were hospitalized between 2003 and 2010 for acute myocardial infarction, hip fracture, gastrointestinal bleeding, stroke, cancer, congestive heart failure, pneumonia, acute respiratory failure, and unstable angina. These diseases comprise a substantial percentage of admissions in the Medicare population and most have been studied previously in the treatment intensity literature.[8,127,140,167] This cohort-based approach is similar to that used in prior studies, including the seminal work of Fisher and colleagues and more recent studies focused on acute myocardial infarction.[4,5,8,127,152,159,167,258]

Healthcare Spending

For each hospitalization, we calculated total Medicare Part A and B spending for the 365-day period beginning with inpatient admission, summing reimbursement for inpatient hospital (including acute care, inpatient rehabilitation, and long-term care), skilled nursing facility, home health or hospice, physician offices, outpatient facility, and durable medical equipment. The measure was price-adjusted using the ratio of price-standardized to price-unstandardized Medicare Part A and B spending within a respondent's hospital referral region (HRR). HRRs represent regional markets for tertiary medical care and define where residents of that region receive the majority of their care.[220] Price-adjustment removes the effects of regional variations in input costs, such as capital, labor, and overhead (rent and liability costs) and more accurately reflects utilization of healthcare services.[124]

Study Outcomes

For each hospitalization, measures of a respondent's pre-hospitalization and post-hospitalization health and functional status were compared using data obtained from their most recent HRS interviews. As such, respondents' outcomes were evaluated at different time periods corresponding to their hospitalizations. On average, interviews were conducted 1.6 years before and 1 year after hospitalization. The latest year of available HRS follow-up data was 2013.

Respondents rated their general health status as poor, fair, good, very good, or excellent. Functional status was measured as a count of limitations based on respondents reports of some difficulty performing basic activities of daily living (ADLs; bathing, eating, dressing, walking across a room, and getting in or out of bed) and instrumental activities of daily living (IADLs; using a telephone, taking medication, handling money, shopping, and preparing meals). Respondents' ratings of pain were classified as none, mild, moderate, or severe. We categorized cognitive functioning as either normal, cognitively impaired without dementia, or with dementia, using a previously validated approach.[261] Depression was measured using an 8-item symptom count adapted for the HRS from the 20-item Center for Epidemiologic Studies Depression (CES-D) scale, which measures depressive symptoms in the general population.[262] Outcome data obtained from proxies representing respondents were included in the main analysis, if available. Additional details about the outcome measures are provided in supplementary Table S4.1 in the Appendix.

We constructed a composite outcome variable that combined information about respondents' pre- versus post-hospitalization scores and their vital status into a single summary measure for each outcome. Respondents scored 0 if their outcome measure was the same or better when compared with pre-hospitalization (denoting benefit), and 1 if their score was worse, or if they had died. This approach sought to mitigate potential bias from truncation-by-death confounding,[215] in which respondents in worse states of health or functioning are less likely to survive to the follow-up HRS interview. For consistency with past research, we also assessed whether each respondent survived to 365 days following hospitalization.

Statistical Analysis

Our study design accounts for two sources of differences in spending across patients: differences due to underlying health, and differences due to treatments received. Because patients with greater severity of illness require more care and incur greater expenditures, failure to account for severity of illness biases estimates toward implying that greater spending is associated with worse health outcomes. We addressed this concern by focusing on an acutely ill patient population and by using instrumental variables analysis—a commonly used econometric technique that can resolve confounding in the key explanatory variable by using a predictor (the "instrument") that is uncorrelated with unobserved health status, has no direct effect on health outcomes (except through its effect on respondent-level spending), and is strongly correlated with the endogenous explanatory variable.[186]

Similar to prior research,[127,179,181] we used HRR-level total price-adjusted Medicare spending per decedent in the last 2 years of life as an instrumental variable; this measure was extracted from *The Dartmouth Atlas of Health Care*.[19] This choice of instrument is supported by a large body of literature demonstrating that significant regional variations in spending and utilization (including at the end of life) persist even after accounting for a comprehensive set of patient characteristics and health status measures.[37,199] Regional end-of-life spending is highly correlated with total spending but less sensitive to differences in illness severity because patients at the end of life are plausibly similar in their health statuses.[5,8,190,263] Physician beliefs and practice patterns rather than patient characteristics drive a significant share of Medicare end-of-life spending.[94]

We first estimated linear probability models for each outcome that did not account for unobserved health status. We then used two-stage least squares to estimate the same models, instrumenting for respondent-level spending with the end-of-life spending from the respondent's HRR of residence. Regressions controlled for demographic and socioeconomic characteristics (age, sex, self-reported race and ethnicity, education, household wealth, and marital status), health behaviors (past and present smoking status, alcohol consumption, and body mass index), a pre-hospitalization measure of the outcome (because changes in health status may differ for respondents with different levels of past health), year of hospitalization, time between outcome assessment (for example, HRS interviews) and admission or discharge, and for the individual diseases comprising the study population. We did not control for comorbidities coded in Medicare claims in order to avoid introducing a bias associated with regional differences in diagnostic intensity.[42] In all analyses, hospitalization was the unit of analysis, and standard errors were clustered by HRR. Additional details of the instrumental variable approach are provided in the supplemental Appendix.

We conducted numerous sensitivity analyses including restricting observations to a respondent's first admission during our study period, excluding hospitalization for congestive heart failure and pneumonia (where there is greater clinical uncertainty around the decision to admit), using alternative specifications of spending and outcome measures, excluding observations where outcome data were obtain by proxy interviews, and clustering standard errors by both HRR and respondents (a complete list of sensitivity analyses is provided in the Appendix).

We used Stata 13.1 (StataCorp, College Station, Texas) for all analyses. Reported P values are two-sided; a P value of 0.05 or less designated statistical significance. This study was approved by the IRB of The Johns Hopkins Bloomberg School of Public Health.

RESULTS

Survey Respondents and Hospitalizations

A total of 8,717 HRS respondents with linked Medicare claims experienced 28,775 hospitalizations between 2003 and 2010. We excluded 2,256 hospitalizations for respondents under age 65, 4,809 hospitalizations for respondents who were not Medicare fee-for-service beneficiaries, and 17,025 hospitalizations that did not meet diagnostic inclusion criteria. The remaining sample included 4,685 hospitalizations for 2,772 respondents residing in 237 HRRs during the study period.

Table 4.1 presents respondent characteristics for all hospitalizations. The mean age was 80.7 years; 56% were women, 23% were nonwhite, and 32% reported fair health status prior to hospitalization. There were few differences in respondent characteristics across HRRs with the following categories of end-of-life spending: low (median, $47,662; range, $34,741 to $52,262), intermediate (median, $59,075; range, $52,339 to $67,422), or high (median, $74,307; range, $67,435 to $110,969). However, notable differences were found in racial and ethnic composition, with a lower percentage of white (63% vs. 85%) and a higher percentage of Hispanic (17% vs. 2%) respondents living in HRRs with high-versus-low end-of-life spending. There was no consistent relationship between pre-hospitalization health and HRR-level spending.

Healthcare Spending

Figure 4.1 compares respondent-level spending in HRRs with low, intermediate, and high end-of-life spending. Respondents' median Medicare spending in the year following hospitalization was $23,858 (interquartile range, $11,664 to $45,476) and was 31% higher when comparing HRRs with high-versus-low end-of-life spending ($29,879 vs. $21,794). HRR-level end-of-life spending—the instrumental variable—was strongly correlated with respondent-level spending (first stage F statistic, 38; P<0.001; Appendix Table S4.2). A $1,000 increase in HRR-level end-of-life spending was associated with a 1.10% (95% confidence interval [CI], 0.75 to 1.50; P<0.001) increase in respondent-level spending (Appendix Table S4.2).

Comparison of Ordinary Least Squares and Instrumental Variables Analyses

Figure 4.2 compares the results when models were estimated by ordinary least squares or instrumental variables. In ordinary least squares regressions that did not account for confounding by severity of illness, we found that higher spending was associated with worse health and functional status. There was no relationship between spending and mortality. In contrast, the instrumental variables analyses, which adjusted for confounding by patient health status, showed beneficial effects of higher spending for certain outcomes, which we discuss in detail hereafter.

Health and Functional Status

In the instrumental variable regressions, a 10% increase in price-adjusted Medicare spending in the year following hospitalization was associated with a reduction in the probability that respondents developed new IADL limitations (-1.96 percentage points; 95% CI, -3.88 to -0.03; P=0.05) and new depressive symptoms (-2.31 percentage points; 95% CI, -4.04 to -0.59; P=0.009) compared with before hospitalization (Table 4.2). These are relative reductions of 3.1% and 3.4% from the mean rates of new IADL limitations and depressive symptoms (63.0% and 68.2%), respectively. There was no significant association between higher Medicare spending and self-rated health status, ADL limitations, pain, or cognitive functioning.

Mortality

Respondents died within 365 days following 37% of hospitalizations. In the instrumental variable regressions, a 10% increase in Medicare spending in the year following hospitalization was associated with a 2.02 percentage point (95% CI, -3.57 to -0.46; P=0.01) reduction in the probability of mortality (Table 4.2), representing a relative decrease of 5.5%.

Sensitivity Analyses

Our findings were robust across numerous sensitivity analyses. Analyses that examined only respondents' first admissions or excluded hospitalizations for congestive heart failure and pneumonia, tested alternative outcome measure specifications, clustered standard errors on respondents rather than HRRs or on both respondents and HRRs, excluded respondents for whom outcome data were obtain by proxy interviews, included Elixhauser comorbidities, or used price-unadjusted measures of spending were not appreciably different from our main findings (Appendix Tables S4.3-S4.5). When comparing hospitalizations in HRRs in the top half of regional end-of-life spending to those in the bottom half, we found that, in higher-intensity regions, a 10% increase in respondent-level spending was associated with reductions in the probability of the following: new IADL limitations and depressive symptoms, worse cognitive status, and 1-year mortality (Appendix Table S4.6). There was no significant association between higher Medicare spending and outcomes in lower-intensity regions. The full results of all sensitivity analyses are provided in the Appendix.

DISCUSSION

Using 10 years' of data from a large, nationally representative study and linked Medicare claims, we examined the effect of healthcare spending on a diverse and understudied set of outcomes that captured beneficiaries' physical, cognitive, and mental health, as well as functional status, and mortality. Higher Medicare spending following hospitalization was associated with minor reductions in the likelihoods of new IADL limitations, new depressive symptoms, and 1-year mortality, but was not associated with other physical and mental health outcomes in the main analysis. Additional analyses demonstrated that in higher-intensity regions, higher spending was also associated with a lower likelihood of cognitive impairment.

This study extends recent research that relied on administrative data, such as discharge and claims data, and found that higher acute care inpatient spending was associated with improved short-term survival.[127,167,258] We instead examined total nondrug healthcare spending in the year following the date of admission. We used this approach because acute care inpatient and post-acute care utilization (skilled nursing, inpatient rehabilitation, long-term care, and home health care) together account for most of the geographic variation in Medicare spending,[35] and because of our interest in longer-term health and functional outcomes that may be particularly responsive to rehabilitative utilization during the post-acute period. Functional disability, cognitive impairment, and depression are highly prevalent among older Americans following hospitalization for serious illnesses,[264-267] though little previous work has assessed whether variations in healthcare spending affect these domains of health. This study underscores the opportunities afforded by large, national surveys such as the HRS to provide estimates of healthcare system productivity and to identify systematic differences in measures of health and functioning that may not be correlated with mortality.

Our instrumental variables results provide evidence of small positive returns to total (acute inpatient and post-acute) spending. While small in magnitude, these reductions in functional disability, cognitive impairment, and depressive symptoms due to greater Medicare spending following hospitalization could offset some of the substantial economic costs associated with informal care—unpaid care provided by family and friends—that many older Americans receive.[268-270] In addition, policies that try to attenuate geographic variation through across-the-board reductions in Medicare spending could unintentionally limit utilization of effective care in addition to care that is wasteful.

This study has several limitations. First, because we studied elderly Medicare fee-for-service beneficiaries, our results may not generalize to Medicare Advantage or commercially insured beneficiaries. Nevertheless, fee-for-service beneficiaries account for over 70% of all Medicare beneficiaries and are the predominant population studied when examining the consequences of variations in healthcare spending.[271] Since our study focused on acutely ill beneficiaries with an inpatient admission, it is unknown whether similar benefits would be achieved by increasing spending within a healthier population. Similarly, because regions have different production functions, it is unclear whether increasing spending within lower-intensity regions would achieve results similar to those observed in higher-intensity regions.[197]

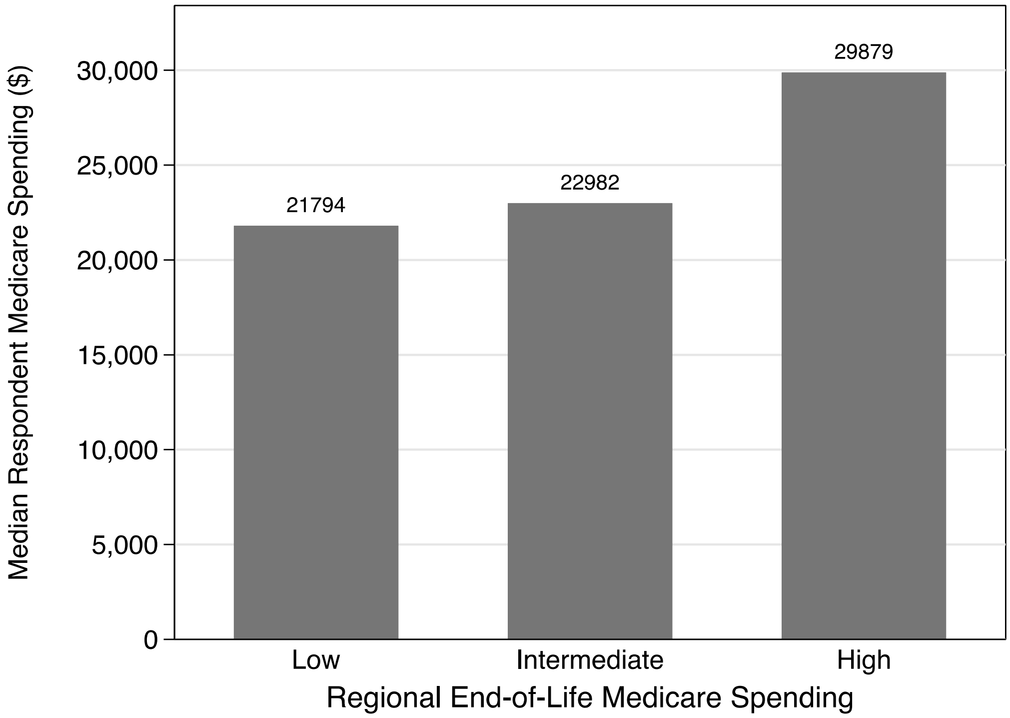
Second, regions with higher intensities of end-of-life spending may have lower thresholds for hospitalizing patients for diseases such as congestive heart failure or pneumonia, whose admission rates—unlike heart attack or hip fracture—are subject to greater provider discretion. This could bias estimates toward showing that lower spending is beneficial, contrary to our main findings. Consistent with other studies,[127,167] our sensitivity analyses suggested that the results were not driven by inclusion of hospitalizations for congestive heart failure or pneumonia; moreover, to adjust for pre-existing differences in health, we included additional covariates (smoking status, alcohol consumption, and body mass index), and pre-hospitalization measures of the outcomes, that are unrelated to the likelihood of hospital admission or regional diagnostic practices. Finally, our study does not disaggregate spending or suggest which sources of spending are beneficial. Certain post-acute spending, such as utilization of skilled nursing facilities, may not be productive for improving patients' survival.[272] Additional research is needed to investigate whether the intensity of utilization of different types of post-acute care services can improve patients' health and functioning.

CONCLUSIONS

We leveraged a large, nationally representative and longitudinal survey of older Americans to examine the relationship between Medicare spending and outcomes that are rarely available in administrative datasets. Higher Medicare spending following hospitalization was associated with fewer new functional limitations and depressive symptoms, in addition to lower mortality rates, but was not related to other measures of health. These findings underscore the importance of assessing a multidimensional set of outcomes as policymakers consider interventions to reduce geographic variation and improve the value of U.S. health care.

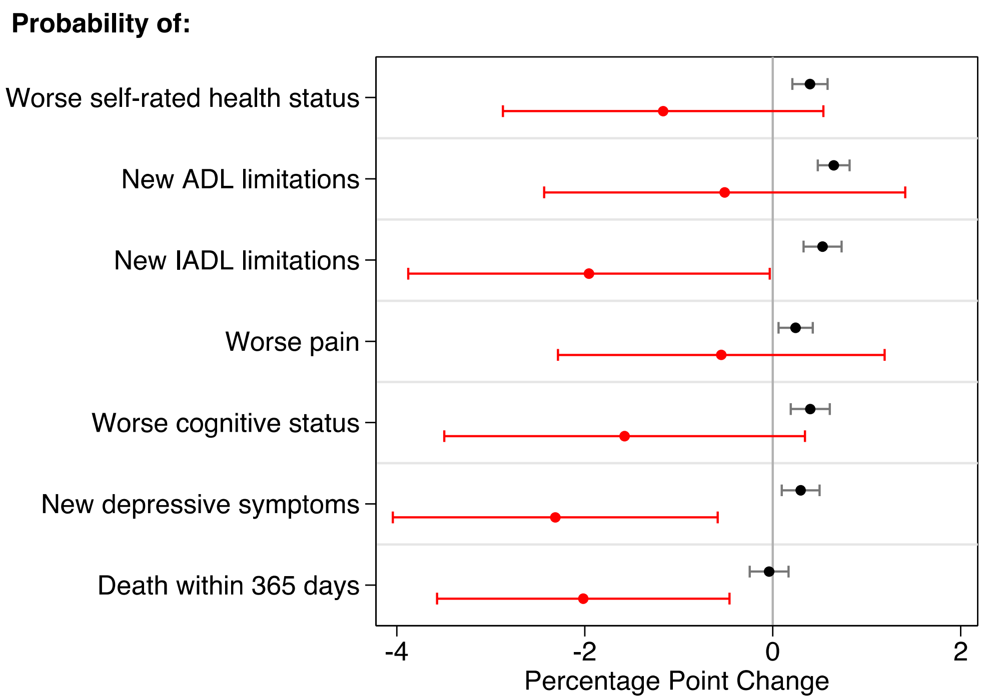
FIGURES

Figure 4.1. Medicare Spending Following Acute Hospitalization



Respondent Medicare spending is the median of total nondrug Medicare Part A and B spending in the year following admission, standardized to adjust for regional differences in price. This includes spending on inpatient hospital care (including acute care, inpatient rehabilitation, and long-term care), skilled nursing facility stays, home health and hospice care, physician services, outpatient care, and durable medical equipment. Regional end-of-life Medicare spending is based on data reported in *The Dartmouth Atlas of Health* Care and refers to total per capita Part A and B spending in the last 2 years of life measured at the hospital referral region (HRR) level and standardized to adjust for regional differences in price. “Low” refers to HRRs equal to or below the 25th percentile of regional end-of-life spending (median, $47,662; range, $34,741 to $52,262), “intermediate” refers to HRRs above the 25th and below the 75th percentiles (median, $59,075; range, $52,339 to $67,422), and “high” refers to HRRs above or equal to the 75th percentile (median, $74,307; range, $67,435 to $110,969).

Figure 4.2. Percentage Point Change in Probability of Outcomes for 10% Increase in Respondents' Medicare Spending Following Hospitalization



Coefficient estimates are from regressions using ordinary least squares (OLS) and two-stage least squares instrumental variable (IV) estimation of the outcomes on respondents' price-adjusted spending and covariates for demographic, health, and time characteristics. All outcomes other than death are measured with respect to health prior to hospitalization; for example, respondents' self-rated health status was worse after versus before hospitalization.

TABLES

Table 4.1. Characteristics of Respondents' Hospitalizations by Regional Intensity of End-of-Life Spending[a]

		Regional End-of-Life Spending			
	All Hospitalizations	Low	Intermediate	High	
Characteristic	**(N = 4,685)**	**(N = 1,176)**	**(N = 2,335)**	**(N = 1,174)**	**P Value**
Age (yr), mean (SD)	80.7 (8.6)	80.6 (8.2)	80.6 (8.7)	81.1 (8.7)	0.71
Female sex (%)	55.5	53.4	56.1	56.2	0.43
Race or ethnic group (%)					
White	77.3	85.0	80.5	63.4	0.002
African American	14.7	12.0	14.8	17.4	0.39
Hispanic	6.3	1.9	3.3	16.7	0.02
Other	1.6	1.1	1.5	2.6	0.32
Education level (%)					
Some high school or less	39.3	38.6	38.5	41.7	0.72
High school diploma	32.8	32.4	33.3	32.3	0.91
More than high school	27.9	29.0	28.3	26.1	0.68
Wealth (%)[b]					
Quintiles 1-3	68.0	66.9	66.8	71.6	0.30
Quintiles 4 and 5	32.0	33.1	33.2	28.4	0.30
Married (%)[c]	53.6	49.5	53.7	57.4	0.05
Self-rated health status, mean (SD)[d]	3.6 (1.1)	3.5 (1.1)	3.6 (1.1)	3.6 (1.1)	0.41
ADL limitations, mean (SD)[e]	1.0 (1.5)	0.9 (1.4)	0.9 (1.5)	1.1 (1.6)	0.006
IADL limitations, mean (SD)[f]	1.0 (1.5)	0.9 (1.4)	1.0 (1.5)	1.1 (1.6)	0.29
Pain (%)					
None	59.9	61.3	58.5	61.3	0.28
Mild	8.6	8.7	8.5	8.8	0.98
Moderate	21.5	19.9	23.0	20.2	0.12
Severe	10.0	10.1	10.0	9.7	0.97
Cognitive status (%)					
Normal	52.1	51.7	54.4	47.8	0.03
Cognitive impairment without dementia	28.8	29.2	27.5	31.1	0.19
Dementia	19.1	19.1	18.0	21.1	0.39
Depressive symptoms, mean (SD)[g]	1.8 (2.1)	1.7 (2.1)	1.8 (2.1)	1.8 (2.1)	0.14

Abbreviations: ADL, activities of daily living; IADL, instrumental activities of daily living.

[a] Age was assessed at the time of hospitalization from respondents' Medicare claims. All other sociodemographic and health characteristics were assessed from respondents' Health and Retirement Study (HRS) surveys prior to hospitalization. Regional end-of-life Medicare spending is based on data reported in *The Dartmouth Atlas of Health* Care and refers to total Part A and B spending in the last 2 years of life, measured at the hospital referral region (HRR) level. "Low" refers to HRRs equal to or below the 25th percentile of regional end-of-life spending (median, $47,662; range, $34,741 to $52,262), "intermediate" refers to HRRs above the 25th and below the 75th percentiles (median, $59,075; range, $52,339 to $67,422), and "high" refers to HRRs above or equal to the 75th percentile (median, $74,307; range, $67,435 to $110,969). P values are from tests of the equality of means across all 3 levels of regional end-of-life spending and account for clustering of hospitalizations within HRRs.

[b] Respondents' total household wealth (excluding individual retirement accounts) was measured in quintiles based on the full wealth distribution of respondents in the corresponding HRS survey wave.

[c] The precise P value was 0.052.

[d] Scores for self-rated health status range from 1 (excellent) to 5 (poor).

[e] Limitations in ADLs are the sum (0 – 5) of difficulties with bathing, eating, dressing, walking across a room, and getting in or out of bed.

[f] Limitations in IADLs are the sum (0 – 5) of difficulties with using a telephone, taking medication, handling money, shopping, and preparing meals.

[g] Depressive symptoms range from 0 to 8 and were assessed using an HRS adaptation of the Center for Epidemiologic Studies Depression scale (CES-D).

Table 4.2. Mean Values and Changes in Measures of Health, Functional Status, and Mortality Following Hospitalization[a]

		Ordinary Least Squares		Instrumental Variables	
Variable	**Mean Value at Follow Up (%)**	**Change with 10% Increase in Spending (95% CI)[b]**	**P Value**	**Change with 10% Increase in Spending (95% CI)[b]**	**P Value**
Measures of health and functional status[c]					
Worse health status	64.9	0.40 (0.21 to 0.58)	<0.001	-1.17 (-2.87 to 0.54)	0.18
New ADL limitations	64.7	0.65 (0.50 to 0.82)	<0.001	-0.51 (-2.43 to 1.41)	0.60
New IADL limitations	63.0	0.53 (0.33 to 0.73)	<0.001	-1.96 (-3.88 to -0.03)	0.05
Worse pain	55.4	0.24 (0.06 to 0.42)	0.009	-0.55 (-2.29 to 1.19)	0.54
Worse cognitive status	58.7	0.40 (0.19 to 0.60)	<0.001	-1.58 (-3.50 to 0.34)	0.11
New depressive symptoms	68.2	0.30 (0.10 to 0.50)	0.004	-2.31 (-4.04 to -0.59)	0.009
Death within 365 days	36.7	-0.04 (-0.24 to 0.17)	0.71	-2.02 (-3.57 to -0.46)	0.01

Abbreviations: ADL, activities of daily living; IADL, instrumental activities of daily living.
[a] The effect of a 10% increase in price-adjusted Medicare spending in the year following hospitalization was estimated from regressions using ordinary least squares or two-stage least squares instrumental variables estimation, with year-specific end-of-life Medicare spending in the respondent's hospital referral region (HRR) as the instrumental variable. Regressions include covariates for age, sex, race and ethnicity, education, household wealth, and marital status; past and present smoking status, alcohol consumption, and body mass index; year of hospitalization; a pre-hospitalization measure of the outcome (using each measure's original scale); the time (in days) between outcome assessment (for example, HRS interviews) and admission or discharge; and for the individual diseases comprising the study population. Standard errors were clustered within HRRs to account for the likelihood that respondents in the same region are treated similarly.
[b] Expressed as the percentage point change in the probability of the outcome (e.g., worse self-rated health status).
[c] Outcomes coded as binary variables with respect to health prior to hospitalization: same or better health = 0 and worse health or death = 1.

APPENDIX

Analytic Approach

We are interested in the relationship between Medicare spending and patient health. However, estimates that fail to correct for unobserved patient severity are likely biased upwards because of a presumed positive correlation between unobserved severity of illness and spending (i.e. higher spending is associated with worse health outcomes). To address this concern, we estimated linear probability models of our composite outcome variables using two-stage least squares (2SLS) instrumental variables regression. In Equation 1, we predict the natural log of Medicare spending using total price-adjusted end-of-life (EOL) Medicare spending measured at the hospital-referral regional (HRR)-level as an instrument. Equation 2 uses the predicted value from the first stage in a regression of patient outcomes on logged one-year Medicare spending. (Y_{i-1} refers to a pre-hospitalization measure of the outcome obtained from the most recent prior Health and Retirement Study interview.)

$$ln\$_i = \beta_0 + \beta_1 \mathrm{HRR\$}_i + \alpha \mathrm{COVARIATES}_i + \delta \mathrm{Y}_{i-1} + \epsilon_i \quad (1)$$

$$\mathrm{Y}_i = \pi_0 + \pi_1 \widehat{ln\$_i} + \alpha \mathrm{COVARIATES}_i + \delta \mathrm{Y}_{i-1} + \xi_i \quad (2)$$

2SLS estimates a local average treatment effect—the effect of healthcare spending for those respondents whose healthcare spending was higher as a result of residing in a region with higher EOL Medicare spending.[273] Hospitalizations were the unit of analysis and standard errors were clustered at the HRR level to account for intraregional correlation.

Defining the Instrumental Variable

We selected the instrumental variable based on theory and past research. In the context of our study, a valid instrumental variable would be exogenous, uncorrelated with unobserved severity of illness, and have no direct effect on outcomes other than through its effect on respondents' spending (the exclusion restriction).[186] Using elements of geography as instrumental variables is common in the empirical health economics and health services literature.[25] For example, Hadley and colleagues (2011),[179] Skinner and colleagues (2005),[181] and Kaestner and colleagues (2010)[127] used measures of regional or hospital-level EOL spending and utilization intensity as instrumental variables.

Regional EOL intensity measures have been used as instruments with the premise that utilization measured in the last 6 months or 2 years of life is unconfounded by severity of illness because all patients have similar health statuses close to the EOL. Kelley and colleagues (2011)[199] found that marked variation in EOL spending remained even after accounting for numerous risk-adjusters derived from HRS survey responses and Medicare claims, suggesting that differences in underlying health did not explain most of this variation. These EOL "look back" spending measures are highly correlated with "look forward" measures that sum the costs of care over a defined period of time for a clinically homogenous cohort.[4,190] As such, EOL spending measures capture a region's proclivity for utilizing health care of varying intensities.

For these reasons, we used total price-adjusted Medicare spending per decedent in the last 2 years of life as an instrumental variable, which we obtained from *The Dartmouth Atlas of Health Care*.[19] The instrument was measured at the HRR level. Details about EOL utilization measures can be found in a Dartmouth Atlas report on chronically ill Medicare beneficiaries at the EOL.[274] In brief, the Dartmouth Atlas constructs this variable by identifying fee-for-service Medicare beneficiaries (with continuous enrollment in Part A and Part B) who were hospitalized with one of nine medical conditions in an acute care hospital in the last 2 years of life. These measures are then adjusted for age, sex, race, primary chronic condition, and the presence of more than one chronic condition. Spending is summed from the following files: Medicare Provider Analysis and Review (MEDPAR), Home Health Agency, Hospice, Durable Medical Equipment, Part B file, and Outpatient.[274]

We used linear interpolation to fill in data for the missing years (2008 and 2009). To implement this measure, we first assigned respondents to 1 of 306 HRRs based on their ZIP code of residence.[19] Next, each hospitalization was assigned a value of the instrumental variable based on the HRR and year of hospitalization. Year-specific information was important because ZIP codes are added and removed over time and the HRRs associated with them also change over time.

Sensitivity Analyses

Appendix Tables S4.3 through S4.6 present the results of the following sensitivity analyses:

A. Only respondents' first (index) admission during the study period included
B. Removal of congestive heart failure (CHF) and pneumonia hospitalizations from the study population
C. Standard errors clustered by respondents rather than HRRs
D. Standard errors clustered by respondents and HRRs
E. Alternative composite outcome measure that only codes deaths following respondents' terminal hospitalizations
F. Outcomes coded on original scales—survivors only
G. Outcomes coded on original scales—survivors and decedents
H. Exclusion of pre-hospitalization measure of the outcome from the model
I. Exclusion of proxy interviews after hospitalization from the study population
J. Price-unadjusted measures of respondent- and HRR-level spending
K. Inclusion of Elixhauser comorbidities
L. Comparison of hospitalizations in high versus low EOL spending HRRs

Table S4.1. Outcome Measure Definitions for Original Scales[a]

Measure	Definition
Self-rated health status	Rated (1 – 5) excellent, very good, good, fair, poor
ADL limitations	Sum (0 – 5) of some difficulties bathing, eating, dressing, walking across a room, and getting in or out of bed
IADL limitations	Sum (0 – 5) of some difficulties with using a telephone, taking medication, handling money, shopping, preparing meals
Cognitive function	Assessed (1 – 3) as normal, CIND, dementia
Pain	Rated (0 – 3) as none, mild, moderate, or severe
Depression (CES-D)[275]	Sum (0 – 8) of affirmative responses to feeling depressed, sad, lonely, everything was an effort, restless sleep, could not get going, and negative responses to feeling happy, and enjoying life

Abbreviations: ADL, activities of daily living; IADL, instrumental activities of daily living; CIND, cognitive impairment without dementia; CES-D, Center for Epidemiologic Studies Depression scale.

Table S4.2. First Stage Regression: Relationship between Respondent Medicare Spending and Regional End-of-Life Spending[a]

Variable	Coefficient (95% CI)
HRR-level per capita Medicare spending ($1,000s) in last two years of life	0.011 (0.0075 to 0.015)
Age less than 70	0.029 (-0.098 to 0.16)
Age 75 - 79	-0.093 (-0.19 to 0.0083)
Age 80 - 84	-0.15 (-0.27 to -0.029)
Age 85 - 89	-0.079 (-0.19 to 0.033)
Age 90 and above	-0.25 (-0.39 to -0.12)
Female	-0.0025 (-0.077 to 0.072)
Black race	0.11 (0.011 to 0.20)
Other race	0.053 (-0.21 to 0.31)
Hispanic	-0.010 (-0.16 to 0.14)
High school diploma	-0.012 (-0.097 to 0.073)
More than high school	-0.0076 (-0.11 to 0.092)
Wealth quintile 2	0.026 (-0.064 to 0.12)
Wealth quintile 3	-0.12 (-0.23 to -0.017)
Wealth quintile 4	-0.090 (-0.20 to 0.020)
Wealth quintile 5	-0.13 (-0.25 to -0.013)
BMI: underweight	-0.014 (-0.16 to 0.13)
BMI: overweight	0.035 (-0.032 to 0.10)
BMI: obese	0.088 (-0.0092 to 0.19)
Drinks alcohol	-0.0070 (-0.077 to 0.063)
Nonsmoker	-0.040 (-0.12 to 0.035)
Current smoker	-0.037 (-0.15 to 0.078)
Divorced/separated	-0.11 (-0.24 to 0.016)
Widowed	-0.025 (-0.11 to 0.056)
Never married	-0.042 (-0.21 to 0.12)
Admitted 2003	0.12 (-0.40 to 0.64)
Admitted 2005	0.062 (-0.038 to 0.16)
Admitted 2006	0.023 (-0.075 to 0.12)
Admitted 2007	0.025 (-0.076 to 0.13)
Admitted 2008	-0.013 (-0.12 to 0.090)
Admitted 2009	0.047 (-0.069 to 0.16)
Admitted 2010	-0.25 (-0.40 to -0.092)
Days since prior interview	0.000069 (-0.000016 to 0.00015)
Days since discharge	0.00022 (0.00011 to 0.00034)
Acute myocardial infarction	0.12 (0.034 to 0.21)
Pneumonia	-0.27 (-0.36 to -0.18)
Cancer	0.021 (-0.11 to 0.16)
Stroke	-0.099 (-0.19 to -0.0026)
Hip fracture	0.42 (0.32 to 0.51)
Gastrointestinal bleeding	-0.39 (-0.56 to -0.22)
Acute respiratory failure	0.11 (-0.028 to 0.24)
Unstable angina	-0.39 (-0.70 to -0.090)
Baseline self-rated health status	0.048 (0.018 to 0.078)
Constant	9.30 (9.06 to 9.53)
Kleibergen-Paap Wald statistic	38.0

Cragg-Donald F statistic	57.5

Abbreviations: HRR, hospital referral region; BMI, body mass index.

[a] A respondent's log-transformed Medicare spending in the year following hospitalization was regressed on HRR-level per capita Medicare spending in last two years of life in the respondent's HRR of residence and other covariates. Both measures of spending were standardized to adjust for regional differences in price. The Cragg-Donald F statistic and Kleibergen-Paap Wald statistic are tests for weak instruments; the Kleibergen-Paap Wald test is a robust version of the Cragg-Donald F statistic. A commonly used rule of thumb proposed by Staiger and Stock (1997)[201] is that an F-statistic below 10 suggests the possibility of a weak instrument problem.

Table S4.3. Sensitivity Analyses A – D, Changes in Measures of Health, Functional Status, and Mortality Associated with a 10% Increase in Medicare Spending[a]

Sensitivity Analysis	Main	A	B	C	D
First-Stage F Statistic[b]	38	24	33	26	37
Sample Size	4,426	2,621	2,129	4,426	4,426
Measures of health and functional status[c]					
Worse health status	-1.17 (0.87)	0.37 (0.98)	-0.0049 (1.10)	-1.17 (0.93)	-1.17 (0.87)
New ADL limitations	-0.51 (0.98)	-0.24 (1.21)	-0.59 (1.09)	-0.51 (0.88)	-0.51 (0.98)
New IADL limitations	-1.96 (0.98)*	-1.79 (1.16)	-1.17 (1.13)	-1.96 (0.97)*	-1.96 (0.98)*
Worse pain	-0.55 (0.89)	0.70 (0.95)	-0.84 (1.14)	-0.55 (0.91)	-0.55 (0.90)
Worse cognitive status	-1.58 (0.98)	0.35 (1.12)	-1.37 (1.27)	-1.58 (1.03)	-1.58 (0.99)
New depressive symptoms	-2.31 (0.88)**	-2.19 (1.33)	-2.14 (1.20)	-2.31 (1.05)*	-2.31 (0.89)**
Death within 365 days	-2.02 (0.79)*	-1.89 (0.96)*	-1.41 (0.80)	-2.02 (0.91)*	-2.02 (0.81)*

Abbreviations: ADL, activities of daily living; IADL, instrumental activities of daily living.

[a] The effect of a 10% increase in price-adjusted Medicare spending in the year following hospitalization was estimated using two-stage least squares instrumental variables estimation, with year-specific end-of-life Medicare spending in the respondent's hospital referral region (HRR) as the instrumental variable. Regressions include covariates for age, sex, race and ethnicity, education, household wealth, and marital status; past and present smoking status, alcohol consumption, and body mass index; year of hospitalization; a pre-hospitalization measure of the outcome; the time (in days) between outcome assessment (for example, HRS interviews) and admission or discharge; and for the individual diseases comprising the study population. Standard errors were clustered within HRRs to account for the likelihood that respondents in the same region are treated similarly. Estimates are expressed as the percentage point change in the probability of the outcome (e.g., worse self-rated health status; standard errors are in parentheses and asterisks denote statistical significance as follows: * for $P < 0.05$, ** for $P < 0.01$, *** for $P < 0.001$.

[b] First stage F statistic is from the Kleibergen-Paap Wald test.

[c] Outcomes coded as binary variables such that same or better health compared with before hospitalization = 0 and worse health or death = 1.

Table S4.4. Sensitivity Analyses E – G, Changes in Measures of Health, Functional Status, and Mortality Associated with a 10% Increase in Medicare Spending[a]

Sensitivity Analysis	**Main[c]**	**E[c]**	**F[d]**	**G[d]**
First-Stage F Statistic[b]	38	28	22	38
Sample Size	4,426	3,341	2,577	4,426
Measures of health and functional status				
Worse health status	-1.17 (0.87)	-1.12 (0.86)	-0.013 (0.026)	-0.037 (0.023)
New ADL limitations	-0.51 (0.98)	-0.36 (1.06)	-0.0012 (0.048)	-0.063 (0.050)
New IADL limitations	-1.96 (0.98)*	-2.13 (1.09)	-0.030 (0.059)	-0.090 (0.054)
Worse pain	-0.55 (0.89)	-0.27 (0.95)	0.034 (0.031)	-0.022 (0.034)
Worse cognitive status	-1.58 (0.98)	-1.60 (1.00)	-0.0078 (0.019)	-0.031 (0.022)
New depressive symptoms	-2.31 (0.88)**	-2.67 (0.97)**	-0.13 (0.056)*	-0.16 (0.063)**
Death within 365 days	-2.02 (0.79)*	-2.02 (0.79)*	-2.02 (0.79)*	-2.02 (0.79)*

Abbreviations: ADL, activities of daily living; IADL, instrumental activities of daily living.

[a] The effect of a 10% increase in price-adjusted Medicare spending in the year following hospitalization was estimated using two-stage least squares instrumental variables estimation, with year-specific end-of-life Medicare spending in the respondent's hospital referral region (HRR) as the instrumental variable. Regressions include covariates for age, sex, race and ethnicity, education, household wealth, and marital status; past and present smoking status, alcohol consumption, and body mass index; year of hospitalization; a pre-hospitalization measure of the outcome; the time (in days) between outcome assessment (for example, HRS interviews) and admission or discharge; and for the individual diseases comprising the study population. Standard errors were clustered within HRRs to account for the likelihood that respondents in the same region are treated similarly. Estimates are expressed as the percentage point change in the probability of the outcome (e.g., worse self-rated health status; standard errors are in parentheses and asterisks denote statistical significance as follows: * for $P < 0.05$, ** for $P < 0.01$, *** for $P < 0.001$.

[b] First stage F statistic is from the Kleibergen-Paap Wald test.

[c] Outcomes coded as binary variables such that same or better health compared with before hospitalization = 0 and worse health or death = 1.

[d] Outcomes coded as scales or counts in these analyses, as in Appendix Table S4.1. For example, in Analysis F, self-rated health is coded as 1=excellent to 5=poor. Sensitivity Analysis F includes only survivors; G includes survivors and decedents with decedents coded to a new level worse (for example, for self-rated health, death would be coded as 6).

Table S4.5. Sensitivity Analyses H – K, Changes in Measures of Health, Functional Status, and Mortality Associated with a 10% Increase in Medicare Spending[a]

Sensitivity Analysis	**Main**	**H**	**I**	**J**	**K**
First-Stage F Statistic[b]	38	40	27	96	38
Sample Size	4,426	4,428	3,897	4,426	4,426
Measures of health and functional status[c]					
Worse health status	-1.17 (0.87)	-1.47 (0.91)	-0.90 (0.87)	-0.83 (0.55)	-1.03 (0.78)
New ADL limitations	-0.51 (0.98)	-0.46 (0.98)	-0.20 (1.02)	-0.70 (0.61)	-0.41 (0.82)
New IADL limitations	-1.96 (0.98)*	-1.91 (0.98)	-2.70 (0.96)**	-1.41 (0.63)*	-1.85 (0.84)*
Worse pain	-0.55 (0.89)	-0.59 (0.88)	-0.26 (0.88)	-0.13 (0.57)	-0.52 (0.77)
Worse cognitive status	-1.58 (0.98)	-1.59 (0.96)	-1.36 (0.96)	-1.01 (0.63)	-1.49 (0.85)
New depressive symptoms	-2.31 (0.88)**	-2.26 (0.97)*	-2.31 (0.88)**	-1.98 (0.55)***	-2.33 (0.84)**
Death within 365 days	-2.02 (0.79)*	-2.02 (0.79)*	-2.51 (0.89)**	-1.50 (0.61)*	-1.89 (0.82)*

Abbreviations: ADL, activities of daily living; IADL, instrumental activities of daily living.

[a] The effect of a 10% increase in price-adjusted Medicare spending in the year following hospitalization was estimated using two-stage least squares instrumental variables estimation, with year-specific end-of-life Medicare spending in the respondent's hospital referral region (HRR) as the instrumental variable. Regressions include covariates for age, sex, race and ethnicity, education, household wealth, and marital status; past and present smoking status, alcohol consumption, and body mass index; year of hospitalization; a pre-hospitalization measure of the outcome; the time (in days) between outcome assessment (for example, HRS interviews) and admission or discharge; and for the individual diseases comprising the study population. Standard errors were clustered within HRRs to account for the likelihood that respondents in the same region are treated similarly. Estimates are expressed as the percentage point change in the probability of the outcome (e.g., worse self-rated health status; standard errors are in parentheses and asterisks denote statistical significance as follows: * for $P < 0.05$, ** for $P < 0.01$, *** for $P < 0.001$.

[b] First stage F statistic is from the Kleibergen-Paap Wald test.

[c] Outcomes coded as binary variables such that same or better health compared with before hospitalization = 0 and worse health or death = 1.

Table S4.6. Comparison of Hospitalizations in High versus Low End-of-Life (EOL) Spending Hospital Referral Regions (Sensitivity Analysis L)[a]

	Ordinary Least Squares		Instrumental Variables	
	Low HRRs	**High HRRs**	**Low HRRs**	**High HRRs**
Sample Size	2,245	2,181	2,245	2,181
First-Stage F Statistic[b]	N/A	N/A	1	39
Measures of health and functional status[c]				
Worse health status	0.34 (0.12)**	0.43 (0.13)**	4.48 (7.16)	-1.98 (1.01)
New ADL limitations	0.71 (0.11)***	0.59 (0.13)	12.2 (10.9)	-1.31 (1.24)
New IADL limitations	0.50 (0.13)***	0.56 (0.16)***	1.75 (4.56)	-2.19 (1.11)*
Worse pain	0.26 (0.13)	0.21 (0.14)	11.8 (11.4)	-1.51 (1.05)
Worse cognitive status	0.38 (0.13)**	0.41 (0.17)*	11.5 (13.9)	-2.35 (1.09)*
New depressive symptoms	0.15 (0.12)	0.45 (0.16)**	12.5 (25.7)	-3.96 (0.91)***
Death within 365 days	-0.19 (0.14)	0.10 (0.15)	-5.40 (5.30)	-2.82 (1.07)**

Abbreviations: ADL, activities of daily living; IADL, instrumental activities of daily living.

[a] Hospitalizations were stratified at the median ($59,063) of regional price-adjusted EOL spending intensity to create high versus low groups. The effect of a 10% increase in price-adjusted Medicare spending in the year following hospitalization was estimated from regressions using ordinary least squares or two-stage least squares instrumental variables estimation, with year-specific end-of-life Medicare spending in the respondent's hospital referral region (HRR) as the instrumental variable. Regressions include covariates for age, sex, race and ethnicity, education, household wealth, and marital status; past and present smoking status, alcohol consumption, and body mass index; year of hospitalization; a pre-hospitalization measure of the outcome; the time (in days) between outcome assessment (for example, HRS interviews) and admission or discharge; and for the individual diseases comprising the study population. Standard errors were clustered within HRRs to account for the likelihood that respondents in the same region are treated similarly. Estimates are expressed as the percentage point change in the probability of the outcome (e.g., worse self-rated health status; standard errors are in parentheses and asterisks denote statistical significance as follows: * for $P < 0.05$, ** for $P < 0.01$, *** for $P < 0.001$.

[b] First stage F statistic is from the Kleibergen-Paap Wald test.

[c] Outcomes coded as binary variables such that same or better health compared with before hospitalization = 0 and worse health or death = 1.

CHAPTER FIVE: DISCUSSION

SUMMARY OF FINDINGS

This dissertation has used nationally representative survey data from the Health and Retirement Study (HRS) linked to Medicare claims to examine (1) whether health behaviors and modifiable risk factors—smoking status, alcohol consumption, body mass index (BMI), and physical activity—contribute to geographic variation in Medicare spending, and (2) the effect of Medicare spending on a diverse and understudied set of outcomes that captured beneficiaries' physical, cognitive, and mental health and functioning, as well as mortality.

Dissertation results demonstrated that in the general fee-for-service Medicare population, behavioral risk factors collectively explained 7% of the difference in spending between higher- and lower-spending regions; therefore, the majority of the difference was not explained by beneficiary characteristics. Among hospitalized beneficiaries, higher Medicare spending following hospitalization was associated with minor reductions in the likelihood of new limitations in instrumental activities of daily living (IADLs), new depressive symptoms, and 1-year mortality. In the main analyses, higher Medicare spending following hospitalization was not associated with other physical and mental health outcomes.

This dissertation provides policymakers with new information about the importance of behavioral risk factors as determinants of regional variation in Medicare spending and the impact that healthcare spending has on multiple dimensions of health and functioning among Medicare beneficiaries.

DISCUSSION OF FINDINGS

Chapter 2 (Manuscript 1) was a review of existing literature on the determinants and consequences of geographic variation in healthcare spending. This existing literature exhibited notable gaps that reflected a lack of studies concerning the potential explanatory role of health behaviors and modifiable risk factors on geographic variation in spending, as well as a paucity of studies examining the consequences of regional spending differences on beneficiaries' health and functional outcomes. Andersen's model of health services utilization was adapted in this dissertation as a conceptual framework that integrated these determinants and outcomes.

To address these gaps in the literature, survey data from the HRS were linked to Medicare claims data and regional spending data from *The Dartmouth Atlas of Health Care*. Together, these data sources achieved broad national geographic coverage by combining data on healthcare utilization (via the claims) with health and functioning outcomes and behaviors (via the survey).[218,219]

Chapter 3 (Manuscript 2) assessed the extent to which beneficiaries' smoking status, alcohol consumption, BMI, and physical activity (among other individual-level characteristics) contributed to regional differences in price-adjusted Medicare spending (the dependent variable).

Chapter 4 (Manuscript 3) evaluated the impact of price-adjusted Medicare spending, as an independent variable, on beneficiaries' health and functional outcomes (the dependent variables) following acute hospitalization between 2003 and 2010 for acute myocardial infarction, hip fracture, gastrointestinal bleeding, stroke, cancer, congestive heart failure, pneumonia, acute respiratory failure, or unstable angina. Because all measures of Medicare spending were price-adjusted, these measures better reflect differences in the utilization of services (rather than differences in price). Importantly, the study populations in Chapters 3 and 4 differed: Chapter 3 included all Medicare beneficiaries, age 65 or older, in fee-for-service Medicare in 2004, whereas Chapter 4 focused on a subset of beneficiaries who similarly were age 65 or older and in fee-for-service Medicare, but were hospitalized for the specific aforementioned diseases between 2003 and 2010. Here, discussion of the main findings from these empiric chapters is reiterated. Additional details can be found in each corresponding chapter.

The results reported in Chapter 3 were largely consistent with prior research. Using data from the Medicare Current Beneficiary Survey (MCBS), Sutherland and colleagues (2009) found that measures of blood pressure, diabetes, BMI, smoking history, and self-rated health explained 18% of the differences in spending between the highest- and lowest-spending regions.[38] Zuckerman and colleagues (2010) also used the MCBS but extended the analysis of Sutherland and colleagues by adding over 10 additional health variables collected in the MCBS—among these, the presence and diagnosis of specific diseases, and whether beneficiaries died or had proxy respondents—and found that health factors explained 29% of the regional differences in spending.[37] Finkelstein and colleagues (2014) identified that patient characteristics may explain a larger share of regional variation in utilization—between 23 and 50%, depending on the analysis—with much of the variation attributable to health factors.[276] However, this study did not directly assess the choices that patients make that influence health (such as health behaviors) and information about health is ultimately still conveyed through claims data and, by extension, physician coding practices. When viewed collectively, the results reported in Chapter 3 and the results of this prior research suggest that beneficiary characteristics account for a nontrivial proportion of regional variation in Medicare spending, but much of the variation remains unexplained by beneficiary characteristics.

However, the results reported in Chapter 3 and this prior literature differ from two recent studies which determined that patient characteristics account for the majority of regional variation in healthcare spending. In the first contradictory study, Reschovsky and colleagues (2013) implemented a modified version of Medicare's Hierarchical Condition Categories model—a risk adjustment model based on billing data—by including only diagnoses they deemed less susceptible to physician discretionary behavior. They found that population health explained between 75% and 85% of spending variations.[39] However, this study, which was based on a sample of physicians rather than on patients, remains susceptible to biases associated with the lower diagnostic thresholds in higher-intensity regions;[42,44] this study may also have over-adjusted regional differences in spending by using the Hierarchical Condition Categories model.[40,41] In order to address the concern of using potentially endogenous measures of health conditions, Chapter 3 included a sensitivity analysis which added the following variables to the regression models used in the decomposition analysis: the self-reported presence of high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, psychiatric condition, or arthritis, as well as an additional variable that summed the number of self-reported health conditions. In total, 16% of the difference in Medicare spending comparing higher-versus-lower-spending regions was explained by beneficiaries' characteristics when this additional set of potentially endogenous variables was included. This result that is virtually identical to the main analysis which lacked these variables (in which 17% was explained by beneficiary characteristics). The lack of sensitivity of the results to these additional variables may be due in part to the use of self-reported health condition measures. Unlike health conditions coded by physicians and captured in Medicare claims data—which are known to be susceptible to biases related to regional variation in physicians' diagnostic practices—these self-reported health condition measures may be less susceptible to such biases. In the second contradictory study, Louise Sheiner (2014) found that 81% of the state-level variation in Medicare spending could be explained by a limited set of aggregate demographic and health measures.[245] By using the state rather than the individual as the unit of analysis, this study is susceptible to the ecological fallacy that the aggregate factors that explain variation at the state level may not be the same factors that explain variation at the patient level.[246] Furthermore, both studies by Reschovsky and Sheiner contradict a diverse literature supporting the importance of provider factors as determinants of geographic variation in spending. Within that literature, past studies have demonstrated that persistent variations in spending exist at the end of life among similarly ill patients,[199] that physicians' preferences and beliefs are geographically correlated and outweigh patients' preferences in driving utilization,[94] and that physicians in higher-spending regions are more likely to recommend discretionary services (e.g., tests of unproven benefit) and to schedule more frequent return visits when compared to physicians in lower-spending regions.[96]

While Chapter 3 focused on beneficiaries' Medicare spending as a dependent variable, Chapter 4 instead used Medicare spending as an independent variable. Therefore, Chapter 4 was predominantly interested in the effects of healthcare utilization on beneficiaries' health and functional outcomes following hospitalization. Among hospitalized beneficiaries, higher Medicare spending following hospitalization was associated with only minor reductions in the likelihood of new IADL limitations, new depressive symptoms, and 1-year mortality. In the main analyses, no associations were observed between higher spending and self-rated health status, limitations in activities of daily living (ADL), pain, or cognitive functioning.

Chapter 4 extended recent research that used discharge and claims data, focusing mainly on the effect of higher acute care inpatient spending on short-term survival, and found that higher acute care inpatient spending was associated with improved survival.[127,167,258] Because the Institute of Medicine's (IOM) 2013 report, *Variation in Health Care Spending: Target Decision Making, Not Geography*, emphasized that acute care inpatient and post-acute care utilization (skilled nursing, inpatient rehabilitation, long-term care, and home health care) together account for most of the geographic variation in Medicare spending,[35] Chapter 4 results instead examined total nondrug healthcare spending in the year following the date of inpatient admission in order to include rehabilitative utilization during the post-acute period that might be particularly impactful on longer-term health and functional outcomes. Similar to the results of other studies in this literature that used instrumental variables estimation, Chapter 4 identified modest but beneficial effects of higher spending.

POLICY IMPLICATIONS

This dissertation has examined topics that are currently of considerable interest to policymakers: application of patient-reported outcomes in research,[277] geographic variation in Medicare spending,[35,123] and the impact of health behaviors on healthcare utilization.[17,82] The specific policy implications of each empiric manuscript are discussed in Chapters 3 and 4, respectively. The focus of this section centers on reconciling the policy implications of the findings from each chapter.

At first blush, the findings of Chapters 3 and 4 appear to have contradictory policy implications—the former suggesting that improvements in beneficiaries' health behaviors and modifiable risk factors could potentially reduce Medicare spending in higher-spending regions, and the latter suggesting that higher spending may be beneficial for certain health and functional outcomes, in which case, implying that reducing spending would be disadvantageous. However, it is necessary to recall that these chapters presented results from different study populations and the results of one chapter should not be generalized to the other. Chapter 3 included general fee-for-service Medicare beneficiaries in 2004, whereas Chapter 4 focused on beneficiaries who were hospitalized for acute episodes of specific diseases. As such, the study population in Chapter 4 was expected, on average, to have worse health status than the general Medicare population. While increased spending within this hospitalized population may yield benefits, it is unknown whether increasing spending in a healthier population (such as the study population used in Chapter 3) would achieve similar benefits. Analogously, Chapter 3 suggests, for example, that increasing rates of physical activity in higher-spending regions may present an opportunity to reduce Medicare spending in these regions through improvements in beneficiaries' health. Such health improvements may reduce the likelihood of hospitalization for the health conditions that comprised the study population in Chapter 4, which would be a benefit to both the Medicare program and its beneficiaries. However, *conditional on being hospitalized*, greater intensity of utilization may nevertheless offer benefits to sicker beneficiaries. The implications of these results should not be conflated.

Findings reported in Chapter 3 also suggest that policies focused on healthcare providers—such as policies that would geographically adjust provider payments—could unfairly reward or penalize providers in certain regions if beneficiary characteristics are not adequately accounted for in regional comparisons. Chapter 4 further suggests that policies that attempt to reduce geographic variation in spending through across-the-board reductions in Medicare spending could unintentionally limit utilization of care that is effective in addition to care that is wasteful. As such, dissertation findings support existing calls for caution in attempting to implement any policies that reduce either aggregate spending or that target higher-spending regions.[3]

STRENGTHS AND LIMITATIONS

This dissertation has several noteworthy strengths. First, the dissertation has leveraged a novel dataset—the HRS linked to Medicare claims—to include a richer set of variables than many previous studies. These variables included the health behaviors and modifiable risk factors of smoking status, alcohol consumption, BMI, and physical activity, as well as household wealth, poverty status, household size, limitations in ADLs and IADLs, pain, cognitive status, and depressive symptoms.

Second, by focusing on behavioral risk factors, as reported in Chapter 3, we sought to isolate the importance of a set of lifestyle characteristics that are both the antecedent contributors to or causes of costly health conditions as well as factors that are potentially mutable. In contrast, prior studies have emphasized beneficiary characteristics regardless of whether those characteristics lend to interventions or modifications that could attenuate spending and improve health. A large body of public health research, and the results of health promotion programs and interventions that can potentially mitigate these behavioral risk factors, underscore the substantial opportunity to reduce morbidity, mortality, and healthcare costs through lifestyle changes.[240-244] Similarly, the results reported in Chapter 4 extend existing literature on the consequences of geographic variation in spending beyond its traditional reliance on mortality. This is vital because much of healthcare utilization does not necessarily impact mortality, but rather affects physical, cognitive, and mental health and functioning. Providing evidence of the benefits to healthcare utilization for these additional outcomes is an important strength and contribution of this dissertation.

Third, this dissertation has used econometric methods in an effort to facilitate causal inference for questions salient to policymakers. Instrumental variables estimation was used to estimate causal effects of healthcare spending and utilization on outcomes, which is an improvement on the observational research in existing literature. Separately, regression decomposition techniques from labor economics were used to quantify the contribution of health behaviors and modifiable risk factors and to explain variations in spending and utilization across regions. With these strengths, this dissertation contributes to the ongoing debate over the determinants and consequences of geographic variation in Medicare spending.

This dissertation, nevertheless, has several limitations. The specific limitations of the empiric manuscripts were discussed in their respective chapters. Here, the discussion centers on the most important limitations for the dissertation as a whole. First, selection bias—largely related to unobserved confounding by severity of illness—is one of the most important threats to validity. The study design of Chapter 4 specifically sought to mitigate this. As discussed in Chapter 2, common approaches to risk adjustment, such as the Elixhauser or Charlson comorbidity schemes, rely on identifying comorbidities present in Medicare claims.[33] These approaches assume that the frequency of diagnosis of comorbidities is independent of intensity of observation, an assumption that several studies have demonstrated is not strictly tenable.[41-44] Appropriate risk adjustment is an area of active research in the geographic variations field.[39] It is especially important when studying elderly patients because medical care may not improve their initial health states but rather attenuate the deterioration of their health status relative to the absence of medical intervention.[179]

In Chapter 4, concerns related to risk adjustment were addressed by: (1) identifying more homogenous disease-specific cohorts that exhibited “low variation” (and therefore were less subject to supply factors or observational intensity because they lend to straightforward diagnoses and consistently result in hospitalization); (2) considering the inclusion of additional covariates derived directly from respondents—(self-reported smoking status, alcohol consumption, and BMI)—variables that were more likely to be free from bias due to regional diagnostic practices compared with comorbidities present in Medicare claims;[41] and (3) using instrumental variables estimation, which could mitigate biases from unmeasured confounders, including severity of illness.

In addition, the decomposition technique used in Chapter 3 did not necessarily estimate causal effects or elucidate underlying mechanistic pathways of behavioral risk factors on spending and utilization, but, nonetheless, provided salient information about the drivers of utilization across regions. Similarly, Chapter 4 reported using aggregate measures of spending, but those did not clarify the causal chain between higher utilization and fewer new functional limitations and depressive symptoms, and lower mortality rates. Additional research is needed to better understand these relationships.

Finally, although fee-for-service Medicare enrollment represents approximately 70% of all Medicare beneficiaries,[278] these dissertation findings may have limited generalizability to patients outside of fee-for-service Medicare, including beneficiaries enrolled in Medicare Advantage plans or privately insured Americans. The IOM report and recent studies have emphasized important differences between Medicare and private insurance markets in terms of the factors that explain geographic variation in healthcare spending.[3,54,279] Whereas differences in the quantity of services utilized appears to explain variation in Medicare spending, price variation is responsible for roughly 70% of geographic variation in spending for privately insured beneficiaries (and quantity of services used is responsible for only 15%).[123]

FUTURE RESEARCH

This dissertation suggests several avenues for future research. First, future studies could take the health and functional outcomes reported in Chapter 4 and apply them to specific clinical populations—such as survivors of critical illnesses who incur long recoveries following hospital discharge—for which the consequences of regional variation are poorly characterized because those data are not captured in administrative datasets. There is wide variation across hospitals in intensive care unit (ICU)-admitting patterns and utilization, including admissions for patients at low risk of death,[280] and this variation is not explained by observable patient or hospital factors.[281] This variation may be emblematic of a lack of professional agreement about how to optimally use the ICU or for which patients it is the appropriate site of care,[282] and, as such, this may signal possible overuse of ICU care. More frequent use of the ICU, particularly in patient populations with little need for it, may increase the risk of iatrogenic complications (such as bloodstream infections and ventilator-associated pneumonia).[281] While one recent study examining an apparently discretionary patient population with pneumonia found that ICU admission was associated with lower mortality and no significant difference in costs,[283] no large-scale studies have estimated causal effects of regional differences in ICU utilization on physical, cognitive, and mental health and functioning. Better understanding of the consequences of variations in ICU utilization is vital to informing physicians' admitting decisions and addressing policy questions about the utility of expanding the ICU bed supply for the aging population.[284] This is especially important given the staggering cost of critical care, which is growing rapidly and is expensive to the U.S. healthcare system,[285] with aggregate costs exceeding 1% of U.S. gross domestic product.[286]

Second, future studies could use econometric methods similar to those used in this dissertation to study the consequences of regional variation in healthcare utilization in private health insurance markets. Third, this dissertation has highlighted the opportunities offered by nationally representative surveys such as the HRS for studying the relevance of beneficiary characteristics—opportunities that are rarely available in administrative datasets alone. Chapter 3, specifically, suggests the potential for improved risk adjustment if health behaviors and modifiable risk factors were collected from other data sources, such as electronic health records. Moreover, future studies can account for these factors when modelling healthcare spending as a function of beneficiary characteristics.

Finally, the results of this dissertation point toward the need in the future to look more closely at provider characteristics—specifically, which services generate health benefits and which provider characteristics contribute to regional variation in spending (and whether those characteristics are acceptable forms of variation). Given the relative importance of non-beneficiary characteristics as emphasized in Chapter 3, possible provider factors for future study may include physicians' education and training, their professional interactions, and local treatment norms.

Viewed collectively, these future lines of inquiry can inform efforts to improve healthcare quality and lower costs in the U.S. Doing so is critical to ensuring the optimal care of Medicare beneficiaries as well as the financial sustainability of the Medicare program.

REFERENCES

1. Glover JA. The incidence of tonsillectomy in school children. *Proc R Soc Med.* 1938;31(10):1219-1236.

2. Wennberg J, Gittelsohn. Small area variations in health care delivery. *Science.* 1973;182(4117):1102-1108.

3. Institute of Medicine Committee on Geographic Variation in Health Care Spending and Promotion of High-Value Care, Newhouse JP, Garber AM, et al. *Variation in health care spending: target decision making, not geography.* Washington, D.C.: The National Academies Press; 2013.

4. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. *Ann Intern Med.* 2003;138(4):288-298.

5. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. *Ann Intern Med.* 2003;138(4):273-287.

6. Cooper RA. States with more health care spending have better-quality health care: lessons about Medicare. *Health Aff (Millwood).* 2009;28(1):w103-115.

7. Skinner J, Chandra A, Goodman D, Fisher ES. The elusive connection between health care spending and quality. *Health Aff (Millwood).* 2009;28(1):w119-123.

8. Ong MK, Mangione CM, Romano PS, et al. Looking forward, looking back: assessing variations in hospital resource use and outcomes for elderly patients with heart failure. *Circ Cardiovasc Qual Outcomes.* 2009;2(6):548-557.

9. Cutler DM. The potential for cost savings in Medicare's future. *Health Aff (Millwood).* 2005;24 Suppl 2:W5R77-80.

10. Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over Medicare reform. *Health Aff (Millwood).* 2002;Suppl Web Exclusives:W96-114.

11. New England Healthcare Institute. Waste and inefficiency in the U.S. health care system. 2008; http://www.nehi.net/writable/publication_files/file/waste_clinical_care_report_final.pdf.

12. McKinsey Global Institute. Accounting for the cost of US health care: a new look at why Americans spend more. 2008;

http://www.mckinsey.com/insights/health_systems_and_services/accounting_for_the_cost_of_us_health_care.

13. Corallo AN, Croxford R, Goodman DC, Bryan EL, Srivastava D, Stukel TA. A systematic review of medical practice variation in OECD countries. *Health Policy.* 2014;114(1):5-14.

14. Centers for Disease Control and Prevention. *The state of aging and health in America 2013.* Atlanta, GA: Centers for Disease Control and Prevention, US Dept of Health and Human Services; 2013.

15. Ford ES, Mokdad AH, Giles WH, Galuska DA, Serdula MK. Geographic variation in the prevalence of obesity, diabetes, and obesity-related behaviors. *Obes Res.* 2005;13(1):118-122.

16. Nguyen K, Marshall L, Hu S, Neff L, Centers for Disease Control and Prevention. State-specific prevalence of current cigarette smoking and smokeless tobacco use among adults aged >/=18 years - United States, 2011-2013. *MMWR Morb Mortal Wkly Rep.* 2015;64(19):532-536.

17. Thorpe KE. The rise in health care spending and what to do about it. *Health Aff (Millwood).* 2005;24(6):1436-1445.

18. Manning WG, Norton EC, Wilk AS. Explaining geographic variation in health care spending, use and quality, and associated methodological challenges. In: Institute of Medicine (U.S.). Committee on Geographic Variation in Health Care Spending and Promotion of High-Value Care, ed. *Variation in health care spending: target decision making, not geography*. Washington DC: National Academies Press; 2012.

19. The Dartmouth Atlas of Health Care. 2014; http://www.dartmouthatlas.org. Accessed March 5, 2015.

20. Nye G. *Geographic variations in health care costs: an exploration of recent studies.* Jayne Koskinas and Ted Giovanis Foundation for Health and Policy; 2014.

21. Paul-Shaheen P, Clark JD, Williams D. Small area analysis: a review and analysis of the North American literature. *J Health Polit Policy Law.* 1987;12(4):741-809.

22. Pauly MV, McGuire TG, Barros PP. *Handbook of health economics.* Amsterdam: North Holland; 2012.

23. Hussey PS, Wertheimer S, Mehrotra A. The association between health care quality and cost: a systematic review. *Ann Intern Med.* 2013;158(1):27-34.

24. Chandra A, Skinner J. Technology growth and expenditure growth in health care. *J Econ Lit.* 2012;50(3):645-680.

25. Skinner J. Causes and consequences of regional variations in health care. In: Pauly MV, McGuire TG, Barros PP, eds. *Handbook of health economics.* Vol 2. Amsterdam: North Holland; 2012:45-93.

26. Wolfe BL. Health status and medical expenditures: is there a link? *Soc Sci Med.* 1986;22(10):993-999.

27. Kachan D, Tannenbaum SL, Olano HA, LeBlanc WG, McClure LA, Lee DJ. Geographical variation in health-related quality of life among older US adults, 1997-2010. *Preventing Chronic Disease.* 2014;11:E110.

28. Fuchs VR. *Who shall live? Health, economics, and social choice.* New York,: Basic Books; 1975.

29. Porell FW, Miltiades HB. Regional differences in functional status among the aged. *Soc Sci Med.* 2002;54(8):1181-1198.

30. Centers for Disease Control and Prevention. State-specific healthy life expectancy at age 65 years--United States, 2007-2009. *MMWR Morb Mortal Wkly Rep.* 2013;62(28):561-566.

31. Pope GC, Kautter J, Ellis RP, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health Care Financ Rev.* 2004;25(4):119-141.

32. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40(5):373-383.

33. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care.* 1998;36(1):8-27.

34. Krause NM, Jay GM. What do global self-rated health items measure? *Med Care.* 1994;32(9):930-942.

35. Newhouse JP, Garber AM. Geographic variation in Medicare services. *N Engl J Med.* 2013;368(16):1465-1468.

36. Congressional Budget Office. *Geographic variation in health care spending.* The Congress of the United States—Congressional Budget Office; February 2008.

37. Zuckerman S, Waidmann T, Berenson R, Hadley J. Clarifying sources of geographic differences in Medicare spending. *N Engl J Med.* 2010;363(1):54-62.

38. Sutherland JM, Fisher ES, Skinner JS. Getting past denial—the high cost of health care in the United States. *N Engl J Med.* 2009;361(13):1227-1230.

39. Reschovsky JD, Hadley J, Romano PS. Geographic variation in fee-for-service Medicare beneficiaries' medical costs is largely explained by disease burden. *Med Care Res Rev.* 2013;70(5):542-563.

40. Skinner J. Comments on James D. Reschovsky et al., “Geographic Variation in Fee-for-Service Medicare Beneficiaries’ Medical Costs Is Largely Explained by Disease Burden,” Medical Care Research & Review, 2013. 2013; http://tdi.dartmouth.edu/press/updates/dartmouth-institute-for-health-policy-and-clincial-practice-responds-to-res. Accessed October 14, 2014.

41. Wennberg DE, Sharp SM, Bevan G, Skinner JS, Gottlieb DJ, Wennberg JE. A population health approach to reducing observational intensity bias in health risk adjustment: cross sectional analysis of insurance claims. *BMJ.* 2014;348:g2392.

42. Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES. Regional variations in diagnostic practices. *N Engl J Med.* 2010;363(1):45-53.

43. Welch HG, Sharp SM, Gottlieb DJ, Skinner JS, Wennberg JE. Geographic variation in diagnosis frequency and risk of death among Medicare beneficiaries. *JAMA.* 2011;305(11):1113-1118.

44. Wennberg JE, Staiger DO, Sharp SM, et al. Observational intensity bias associated with illness adjustment: cross sectional analysis of insurance claims. *BMJ.* 2013;346:f549.

45. Pine M, Jordan HS, Elixhauser A, et al. Enhancement of claims data to improve risk adjustment of hospital mortality. *JAMA.* 2007;297(1):71-76.

46. Bernstein AB, Hing E, Moss AJ, Allen KF, Siller AB, Tiggle RB. *Health care in America: trends in utilization.* Hyattsville, Maryland: National Center for Health Statistics; 2003.

47. Hibbard JH, Pope CR. Gender roles, illness orientation and use of medical services. *Soc Sci Med.* 1983;17(3):129-137.

48. Bertakis KD, Azari R, Helms LJ, Callahan EJ, Robbins JA. Gender differences in the utilization of health care services. *J Fam Pract.* 2000;49(2):147-152.

49. Mustard CA, Kaufert P, Kozyrskyj A, Mayer T. Sex differences in the use of health care services. *N Engl J Med.* 1998;338(23):1678-1683.

50. Institute of Medicine. *Unequal treatment: confronting racial and ethnic disparities in health care.* Washington DC: National Academies Press; 2003.

51. Baicker K, Chandra A, Skinner JS, Wennberg JE. Who you are and where you live: how race and geography affect the treatment of medicare beneficiaries. *Health Aff (Millwood).* 2004;Suppl Variation:VAR33-44.

52. Baicker K, Chandra A, Skinner JS. Geographic variation in health care and the problem of measuring racial disparities. *Perspect Biol Med.* 2005;48(1 Suppl):S42-53.

53. Chandra A, Skinner J. Geography and racial health disparities. *National Bureau of Economic Research Working Paper Series.* 2003;No. 9513.

54. Gottlieb DJ, Zhou W, Song Y, Andrews KG, Skinner JS, Sutherland JM. Prices don't drive regional Medicare spending variations. *Health Aff (Millwood).* 2010;29(3):537-543.

55. Fiscella K, Franks P, Gold MR, Clancy CM. Inequality in quality: addressing socioeconomic, racial, and ethnic disparities in health care. *JAMA.* 2000;283(19):2579-2584.

56. Adler NE, Newman K. Socioeconomic disparities in health: pathways and policies. *Health Aff (Millwood).* 2002;21(2):60-76.

57. Ross CE, Wu C-l. The links between education and health. *American Sociological Review.* 1995;60(5):719-745.

58. Hadley J. Sicker and poorer—the consequences of being uninsured: a review of the research on the relationship between health insurance, medical care use, health, work, and income. *Med Care Res Rev.* 2003;60(2 Suppl):3S-75S; discussion 76S-112S.

59. Pollack CE, Chideya S, Cubbin C, Williams B, Dekker M, Braveman P. Should health studies measure wealth? A systematic review. *Am J Prev Med.* 2007;33(3):250-264.

60. Allin S, Masseria C, Mossialos E. Measuring socioeconomic differences in use of health care services by wealth versus by income. *Am J Public Health.* 2009;99(10):1849-1855.

61. Angell M. Privilege and health—what is the connection? *N Engl J Med.* 1993;329(2):126-127.

62. Baicker K, Taubman SL, Allen HL, et al. The Oregon experiment—effects of Medicaid on clinical outcomes. *N Engl J Med.* 2013;368(18):1713-1722.

63. Sturm R. The effects of obesity, smoking, and drinking on medical problems and costs. *Health Aff (Millwood).* 2002;21(2):245-253.

64. Bertakis KD, Azari R. The influence of obesity, alcohol abuse, and smoking on utilization of health care services. *Fam Med.* 2006;38(6):427-434.

65. Alexandre PK, Roebuck MC, French MT, Chitwood DD, McCoy CB. Problem drinking, health services utilization, and the cost of medical care. *Recent Dev Alcohol.* 2001;15:285-298.

66. Bertakis KD, Azari R. Obesity and the use of health care services. *Obes Res.* 2005;13(2):372-379.

67. Fries JF. Measuring and monitoring success in compressing morbidity. *Ann Intern Med.* 2003;139(5 Pt 2):455-459.

68. de Rezende LF, Rey-Lopez JP, Matsudo VK, do Carmo Luiz O. Sedentary behavior and health outcomes among older adults: a systematic review. *BMC Public Health.* 2014;14:333.

69. Koplan JP, Dietz WH. Caloric imbalance and public health policy. *JAMA.* 1999;282(16):1579-1581.

70. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. *Br Med J.* 1950;2(4682):739-748.

71. Center for Disease Control and Prevention. *How tobacco smoke causes disease: the biology and behavioral basis for smoking-attributable disease: a report of the Surgeon General.* Atlanta, GA: Publications and Reports of the Surgeon General; 2010.

72. U.S. Department of Health and Human Services. *The health consequences of smoking—50 years of progress: a report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014.

73. Ronksley PE, Brien SE, Turner BJ, Mukamal KJ, Ghali WA. Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. *BMJ.* 2011;342:d671.

74. Mukamal KJ, Conigrave KM, Mittleman MA, et al. Roles of drinking pattern and type of alcohol consumed in coronary heart disease in men. *N Engl J Med.* 2003;348(2):109-118.

75. Chokshi DA, El-Sayed AM, Stine NW. J-shaped curves and public health. *JAMA.* 2015:1-3.

76. Reynolds K, Lewis B, Nolen JD, Kinney GL, Sathya B, He J. Alcohol consumption and risk of stroke: a meta-analysis. *JAMA.* 2003;289(5):579-588.

77. Rehm J, Room R, Graham K, Monteiro M, Gmel G, Sempos CT. The relationship of average volume of alcohol consumption and patterns of drinking to burden of disease: an overview. *Addiction.* 2003;98(9):1209-1228.

78. Bagnardi V, Blangiardo M, La Vecchia C, Corrao G. A meta-analysis of alcohol drinking and cancer risk. *Br J Cancer.* 2001;85(11):1700-1705.

79. Corrao G, Bagnardi V, Zambon A, La Vecchia C. A meta-analysis of alcohol consumption and the risk of 15 diseases. *Prev Med.* 2004;38(5):613-619.

80. Rehm J, Gmel G, Sempos CT, Trevisan M. Alcohol-related morbidity and mortality. *Alcohol Res Health.* 2003;27(1):39-51.

81. Field AE, Coakley EH, Must A, et al. Impact of overweight on the risk of developing common chronic diseases during a 10-year period. *Arch Intern Med.* 2001;161(13):1581-1586.

82. Thorpe KE, Ogden LL, Galactionova K. Chronic conditions account for rise in Medicare spending from 1987 to 2006. *Health Aff (Millwood).* 2010;29(4):718-724.

83. Cawley J, Meyerhoefer C. The medical care costs of obesity: an instrumental variables approach. *J Health Econ.* 2012;31(1):219-230.

84. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. *JAMA.* 2004;291(10):1238-1245.

85. McGinnis JM, Foege WH. Actual causes of death in the United States. *JAMA.* 1993;270(18):2207-2212.

86. Room R. Measuring alcohol consumption in the United States: methods and rationales. In: Kozlowski LT, Annis HM, Cappell HD, et al., eds. *Research advances in alcohol and drug problems.* Vol 10. New York: Plenum Press; 1990.

87. Ewing JA. Detecting alcoholism. The CAGE questionnaire. *JAMA.* 1984;252(14):1905-1907.

88. Riekert KA, Ockene JK, Pbert L. *Handbook of health behavior change.* 4th ed. New York, NY: Springer Publishing Company; 2014.

89. Anthony DL, Herndon MB, Gallagher PM, et al. How much do patients' preferences contribute to resource use? *Health Aff (Millwood).* 2009;28(3):864-873.

90. Wennberg JE. *Tracking medicine: a researcher's quest to understand health care.* New York: Oxford University Press; 2010.

91. Fisher ES, Wennberg JE. Health care quality, geographic variations, and the challenge of supply-sensitive care. *Perspect Biol Med.* 2003;46(1):69-79.

92. Baker LC, Bundorf MK, Kessler DP. Patients' preferences explain a small but significant share of regional variation In Medicare spending. *Health Affairs.* 2014;33(6):957-963.

93. Barnato AE, Herndon MB, Anthony DL, et al. Are regional variations in end-of-life care intensity explained by patient preferences? A study of the US Medicare population. *Med Care.* 2007;45(5):386-393.

94. Cutler D, Skinner J, Stern AD, Wennberg D. Physician beliefs and patient preferences: a new look at regional variation in health care spending. *National Bureau of Economic Research Working Paper Series.* 2013;No. 13301.

95. Eisenberg JM. Physician utilization: the state of research about physicians' practice patterns. *Med Care.* 2002;40(11):1016-1035.

96. Sirovich B, Gallagher PM, Wennberg DE, Fisher ES. Discretionary decision making by primary care physicians and the cost of U.S. Health care. *Health Aff (Millwood).* 2008;27(3):813-823.

97. Gerrity MS, DeVellis RF, Earp JA. Physicians' reactions to uncertainty in patient care. A new measure and new insights. *Med Care.* 1990;28(8):724-736.

98. Molitor D. The evolution of physician practice styles: evidence from cardiologist migration. 2011. http://economics.mit.edu/files/7301.

99. Ghosh AK. On the challenges of using evidence-based information: the role of clinical uncertainty. *J Lab Clin Med.* 2004;144(2):60-64.

100. Wennberg JE, Barnes BA, Zubkoff M. Professional uncertainty and the problem of supplier-induced demand. *Soc Sci Med.* 1982;16(7):811-824.

101. Burke MA, Fournier GM, Prasad K. Geographic variations in a model of physician treatment choice with social interactions. *J Econ Behav Organ.* 2010;73(3):418-432.

102. de Jong JD, Westert GP, Lagoe R, Groenewegen PP. Variation in hospital length of stay: do physicians adapt their length of stay decisions to what is usual in the hospital where they work? *Health Serv Res.* 2006;41(2):374-394.

103. de Jong JD, Groenewegen PP, Westert GP. Mutual influences of general practitioners in partnerships. *Soc Sci Med.* 2003;57(8):1515-1524.

104. Tamblyn R, McLeod P, Hanley JA, Girard N, Hurley J. Physician and practice characteristics associated with the early utilization of new prescription drugs. *Med Care.* 2003;41(8):895-908.

105. Epstein AJ, Nicholson S. The formation and evolution of physician treatment styles: an application to cesarean sections. *J Health Econ.* 2009;28(6):1126-1140.

106. Goodrick E, Salancik GR. Organizational discretion in responding to institutional practices: hospitals and Cesarean births. *Administrative Science Quarterly.* 1996;41(1):1-28.

107. de Jong JD, Groenewegen PP, Spreeuwenberg P, Schellevis F, Westert GP. Do guidelines create uniformity in medical practice? *Soc Sci Med.* 2010;70(2):209-216.

108. Reames BN, Shubeck SP, Birkmeyer JD. Strategies for reducing regional variation in the use of surgery: a systematic review. *Ann Surg.* 2014;259(4):616-627.

109. Arrow KJ. Uncertainty and the welfare economics of medical care. *American Economic Review.* 1963;53:941-973.

110. Newhouse JP. Toward a theory of nonprofit institutions: an economic model of a hospital. *American Economic Review.* 1970;60(1):64-74.

111. Sloan FA, Picone GA, Taylor DH, Chou SY. Hospital ownership and cost and quality of care: is there a dime's worth of difference? *J Health Econ.* 2001;20(1):1-21.

112. Roemer MI. Bed supply and hospital utilization: a natural experiment. *Hospitals.* 1961;35:36-42.

113. Cutler DM, Sheiner L. The geography of Medicare. *American Economic Review.* 1999;89(2):228-233.

114. Delamater PL, Messina JP, Grady SC, WinklerPrins V, Shortridge AM. Do more hospital beds lead to higher hospitalization rates? a spatial examination of Roemer's Law. *PLoS One.* 2013;8(2):e54900.

115. Knickman JR, Foltz AM. A statistical analysis of reasons for East-West differences in hospital use. *Inquiry.* 1985;22(1):45-58.

116. Wennberg JE, Freeman JL, Shelton RM, Bubolz TA. Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. *N Engl J Med.* 1989;321(17):1168-1173.

117. American Hospital Association. *Geographic variation in health care spending: a closer look.* Washington DC: TrendWatch; 2009.

118. Wolfe BL, Detmer DE. The economics of surgical signatures. *Hosp Med Staff.* 1984;13(10):2-8.

119. Baicker K, Chandra A. Medicare spending, the physician workforce, and beneficiaries' quality of care. *Health Aff (Millwood).* 2004;Suppl Web Exclusives:W4-184-197.

120. Wennberg JE, Fisher ES, Stukel TA, Skinner JS, Sharp SM, Bronner KK. Use of hospitals, physician visits, and hospice care during last six months of life among cohorts loyal to highly respected hospitals in the United States. *BMJ.* 2004;328(7440):607.

121. Fisher ES, Wennberg JE, Stukel TA, Sharp SM. Hospital readmission rates for cohorts of Medicare beneficiaries in Boston and New Haven. *N Engl J Med.* 1994;331(15):989-995.

122. Grytten J, Sorensen R. Practice variation and physician-specific effects. *J Health Econ.* 2003;22(3):403-418.

123. Newhouse JP, Garber AM. Geographic variation in health care spending in the United States: insights from an Institute of Medicine report. *JAMA.* 2013;310(12):1227-1228.

124. Gold M. Geographic variation in Medicare per capita spending: should policy-makers be concerned? *The synthesis project: new insights from research results.* Princeton, NJ: Robert Wood Johnson Foundation; 2004.

125. Bradbury RC, Golec JH, Steen PM. Relating hospital health outcomes and resource expenditures. *Inquiry.* 1994;31(1):56-65.

126. Bradbury RC, Golec JH, Steen PM. Toward a systems quality paradigm: relating health outcomes, resource expenditures, and appropriateness of cholecystectomy patients. *Health Serv Manage Res.* 1997;10(4):231-244.

127. Kaestner R, Silber JH. Evidence on the efficacy of inpatient spending on Medicare patients. *Milbank Q.* 2010;88(4):560-594.

128. Bradbury RC, Golec JH, Steen PM. Linking health outcomes and resource efficiency for hospitalized patients: do physicians with low mortality and morbidity rates also have low resource expenditures? *Health Serv Manage Res.* 2000;13(1):57-68.

129. Auerbach AD, Hilton JF, Maselli J, Pekow PS, Rothberg MB, Lindenauer PK. Case volume, quality of care, and care efficiency in coronary artery bypass surgery. *Arch Intern Med.* 2010;170(14):1202-1208.

130. Carey K, Burgess JF, Jr. On measuring the hospital cost/quality trade-off. *Health Econ.* 1999;8(6):509-520.

131. Chen LM, Jha AK, Guterman S, Ridgway AB, Orav EJ, Epstein AM. Hospital cost of care, quality of care, and readmission rates: penny wise and pound foolish? *Arch Intern Med.* 2010;170(4):340-346.

132. Deily ME, McKay NL. Cost inefficiency and mortality rates in Florida hospitals. *Health Econ.* 2006;15(4):419-431.

133. Fleming ST. The relationship between quality and cost: pure and simple? *Inquiry.* 1991;28(1):29-38.

134. Glance LG, Dick AW, Osler TM, Meredith W, Mukamel DB. The association between cost and quality in trauma: is greater spending associated with higher-quality care? *Ann Surg.* 2010;252(2):217-222.

135. Huerta TR, Ford EW, Peterson LT, Brigham KH. Testing the hospital value proposition: an empirical analysis of efficiency and quality. *Health Care Manage Rev.* 2008;33(4):341-349.

136. Jha AK, Orav EJ, Dobson A, Book RA, Epstein AM. Measuring efficiency: the association of hospital costs and quality of care. *Health Aff (Millwood).* 2009;28(3):897-906.

137. Lagu T, Rothberg MB, Nathanson BH, Pekow PS, Steingrub JS, Lindenauer PK. The relationship between hospital spending and mortality in patients with sepsis. *Arch Intern Med.* 2011;171(4):292-299.

138. McKay NL, Deily ME. Cost inefficiency and hospital health outcomes. *Health Econ.* 2008;17(7):833-848.

139. Morey RC, Fine DJ, Loree SW, Retzlaff-Roberts DL, Tsubakitani S. The trade-off between hospital cost and quality of care. An exploratory empirical analysis. *Med Care.* 1992;30(8):677-698.

140. Mukamel DB, Zwanziger J, Bamezai A. Hospital competition, resource allocation and quality of care. *BMC Health Serv Res.* 2002;2(1):10.

141. Mukamel DB, Zwanziger J, Tomaszewski KJ. HMO penetration, competition, and risk-adjusted hospital mortality. *Health Serv Res.* 2001;36(6 Pt 1):1019-1035.

142. Picone GA, Sloan FA, Chou S-Y, Taylor DH, Jr. Does higher hospital cost imply higher quality of care? *Review of Economics and Statistics.* 2003;85(1):51-62.

143. Romley JA, Goldman DP. How costly is hospital quality? A revealed-preference approach. *J Ind Econ.* 2011;59(4):578-608.

144. Saleh S, Callan M, Kassak K. The association between the hospital quality alliance's pneumonia measures and discharge costs. *J Health Care Finance.* 2012;38(3):50-60.

145. Anderson RA, Hsieh PC, Su HF. Resource allocation and resident outcomes in nursing homes: comparisons between the best and worst. *Res Nurs Health.* 1998;21(4):297-313.

146. Hicks LL, Rantz MJ, Petroski GF, Mukamel DB. Nursing home costs and quality of care outcomes. *Nurs Econ.* 2004;22(4):178-192, 175.

147. Mukamel DB, Spector WD. Nursing home costs and risk-adjusted outcome measures of quality. *Med Care.* 2000;38(1):78-89.

148. Weech-Maldonado R, Shea D, Mor V. The relationship between quality of care and costs in nursing homes. *Am J Med Qual.* 2006;21(1):40-48.

149. Weech-Maldonado R, Neff G, Mor V. Does quality of care lead to better financial performance?: the case of the nursing home industry. *Health Care Manage Rev.* 2003;28(3):201-216.

150. Doyle JJ, Jr., Graves JA, Gruber J, Kleiner S. Do high-cost hospitals deliver better care? Evidence from ambulance referral patterns. *National Bureau of Economic Research Working Paper Series.* 2012;No. 17936.

151. Baicker K, Buckles KS, Chandra A. Geographic variation in the appropriate use of cesarean delivery. *Health Aff (Millwood).* 2006;25(5):w355-367.

152. Doyle JJ. Returns to local-area health care spending: using health shocks to patients far from home. *National Bureau of Economic Research Working Paper Series.* 2007;No. 13301.

153. Landrum MB, Meara ER, Chandra A, Guadagnoli E, Keating NL. Is spending more always wasteful? The appropriateness of care and outcomes among colorectal cancer patients. *Health Aff (Millwood).* 2008;27(1):159-168.

154. Sirovich BE, Gottlieb DJ, Welch HG, Fisher ES. Regional variations in health care intensity and physician perceptions of quality of care. *Ann Intern Med.* 2006;144(9):641-649.

155. Wennberg JE, Bronner K, Skinner JS, Fisher ES, Goodman DC. Inpatient care intensity and patients' ratings of their hospital experiences. *Health Aff (Millwood).* 2009;28(1):103-112.

156. Yasaitis L, Fisher E, Mackenzie TA, Wasson J. Healthcare intensity is associated with lower ratings of healthcare quality by younger adults. *J Ambul Care Manage.* 2009;32(3):226-231.

157. Mangione CM, Gerzoff RB, Williamson DF, et al. The association between quality of care and the intensity of diabetes disease management programs. *Ann Intern Med.* 2006;145(2):107-116.

158. Barnato AE, Chang CC, Farrell MH, Lave JR, Roberts MS, Angus DC. Is survival better at hospitals with higher "end-of-life" treatment intensity? *Med Care.* 2010;48(2):125-132.

159. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ. Variations in the longitudinal efficiency of academic medical centers. *Health Aff (Millwood).* 2004;Suppl Variation:VAR19-32.

160. Silber JH, Kaestner R, Even-Shoshan O, Wang Y, Bressler LJ. Aggressive treatment style and surgical outcomes. *Health Serv Res.* 2010;45(6 Pt 2):1872-1892.

161. Yasaitis L, Fisher ES, Skinner JS, Chandra A. Hospital quality and intensity of spending: is there an association? *Health Aff (Millwood).* 2009;28(4):w566-572.

162. Byrne MM, Pietz K, Woodard L, Petersen LA. Health care funding levels and patient outcomes: a national study. *Health Econ.* 2007;16(4):385-393.

163. Fowler FJ, Jr., Gallagher PM, Anthony DL, Larsen K, Skinner JS. Relationship between regional per capita Medicare expenditures and patient perceptions of quality of care. *JAMA.* 2008;299(20):2406-2412.

164. Roski J, Turbyville S, Dunn D, Krushat M, Scholle SH. Resource use and associated care effectiveness results for people with diabetes in managed care organizations. *Am J Med Qual.* 2008;23(5):365-374.

165. Birkmeyer JD, Gust C, Dimick JB, Birkmeyer NJ, Skinner JS. Hospital quality and the cost of inpatient surgery in the United States. *Ann Surg.* 2012;255(1):1-5.

166. Englesbe MJ, Dimick JB, Fan Z, Baser O, Birkmeyer JD. Case mix, quality and high-cost kidney transplant patients. *Am J Transplant.* 2009;9(5):1108-1114.

167. Romley JA, Jena AB, Goldman DP. Hospital spending and inpatient mortality: evidence from California: an observational study. *Ann Intern Med.* 2011;154(3):160-167.

168. Zhang B, Wright AA, Huskamp HA, et al. Health care costs in the last week of life: associations with end-of-life conversations. *Arch Intern Med.* 2009;169(5):480-488.

169. Grabowski DC. Does an increase in the Medicaid reimbursement rate improve nursing home quality? *J Gerontol B Psychol Sci Soc Sci.* 2001;56(2):S84-93.

170. Cohen JW, Spector WD. The effect of Medicaid reimbursement on quality of care in nursing homes. *J Health Econ.* 1996;15(1):23-48.

171. Grabowski DC. Medicaid reimbursement and the quality of nursing home care. *J Health Econ.* 2001;20(4):549-569.

172. Rosenthal MB, de Brantes FS, Sinaiko AD, Frankel M, Robbins RD, Young S. Bridges to Excellence—recognizing high-quality care: analysis of physician quality and resource use. *Am J Manag Care.* 2008;14(10):670-677.

173. Starfield B, Powe NR, Weiner JR, et al. Costs vs quality in different types of primary care settings. *JAMA.* 1994;272(24):1903-1908.

174. Kralewski JE, Dowd BE, Xu Y. Differences in the cost of health care provided by group practices in Minnesota. *Minn Med.* 2011;94(2):41-44.

175. Solberg LI, Lyles CA, Shore AD, Lemke KW, Weiner JP. Is quality free? The relationship between cost and quality across 18 provider groups. *Am J Manag Care.* 2002;8(5):413-422.

176. Cunningham PJ. High medical cost burdens, patient trust, and perceived quality of care. *J Gen Intern Med.* 2009;24(3):415-420.

177. Fenton JJ, Jerant AF, Bertakis KD, Franks P. The cost of satisfaction: a national study of patient satisfaction, health care utilization, expenditures, and mortality. *Arch Intern Med.* 2012;172(5):405-411.

178. Fu AZ, Wang N. Healthcare expenditures and patient satisfaction: cost and quality from the consumer's perspective in the US. *Curr Med Res Opin.* 2008;24(5):1385-1394.

179. Hadley J, Waidmann T, Zuckerman S, Berenson RA. Medical spending and the health of the elderly. *Health Serv Res.* 2011;46(5):1333-1361.

180. Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q.* 2005;83(4):691-729.

181. Skinner J, Fisher ES, Wennberg J. The efficiency of Medicare. In: Wise DA, ed. *Analyses in the Economics of Aging*: University of Chicago Press; 2005.

182. Erickson P. Evaluation of a population-based measure of quality of life: the Health and Activity Limitation Index (HALex). *Qual Life Res.* 1998;7(2):101-114.

183. Cameron AC, Trivedi PK. *Microeconometrics: methods and applications.* Cambridge: Cambridge University Press; 2005.

184. Dunning T. *Natural experiments in the social sciences: a design-based approach.* Cambridge: Cambridge University Press; 2012.

185. Angrist JD, Imbens GW, Rubin DB. Identification of Causal Effects Using Instrumental Variables. *JASA.* 1996;91(434):444-455.

186. Newhouse JP, McClellan M. Econometrics in outcomes research: the use of instrumental variables. *Annu Rev Public Health.* 1998;19:17-34.

187. Ali MS, Groenwold RH, Klungel OH. Propensity score methods and unobserved covariate imbalance: comments on "squeezing the balloon". *Health Serv Res.* 2014;49(3):1074-1082.

188. Stukel TA, Fisher ES, Alter DA, et al. Association of hospital spending intensity with mortality and readmission rates in Ontario hospitals. *JAMA.* 2012;307(10):1037-1045.

189. O'Hare AM, Rodriguez RA, Hailpern SM, Larson EB, Kurella Tamura M. Regional variation in health care intensity and treatment practices for end-stage renal disease in older adults. *JAMA.* 2010;304(2):180-186.

190. Bach PB. A map to bad policy—hospital efficiency measures in the Dartmouth Atlas. *N Engl J Med.* 2010;362(7):569-573.

191. Bach PB, Schrag D, Begg CB. Resurrecting treatment histories of dead patients: a study design that should be laid to rest. *JAMA.* 2004;292(22):2765-2770.

192. Stukel TA, Lucas FL, Wennberg DE. Long-term outcomes of regional variations in intensity of invasive vs medical management of Medicare Patients with acute myocardial infarction. *JAMA.* 2005;293(11):1329-1337.

193. Wooldridge JM. *Introductory econometrics: a modern approach.* 5th ed. Mason, OH: South-Western Cengage Learning; 2013.

194. Elwert F, Winship C. Endogenous selection bias: the problem of conditioning on a collider variable. *Annual Review of Sociology.* 2014;40(1):31-53.

195. Schreyogg J, Stargardt T. The trade-off between costs and outcomes: the case of acute myocardial infarction. *Health Serv Res.* 2010;45(6 Pt 1):1585-1601.

196. Card D, Dobkin C, Maestas N. The impact of nearly universal insurance coverage on health care utilization: evidence from Medicare. *Am Econ Rev.* 2008;98(5):2242-2258.

197. Chandra A, Staiger DO. Productivity spillovers in healthcare: evidence from the treatment of heart attacks. *J Polit Econ.* 2007;115:103-140.

198. Doyle JJ, Jr. Health insurance, treatment and outcomes: using auto accidents as health shocks. *Review of Economics and Statistics.* 2005;87(2):256-270.

199. Kelley AS, Ettner SL, Morrison RS, Du Q, Wenger NS, Sarkisian CA. Determinants of medical expenditures in the last 6 months of life. *Ann Intern Med.* 2011;154(4):235-242.

200. Grootendorst P. A review of instrumental variables estimation of treatment effects in the applied health sciences. *Health Serv Outcomes Res Method.* 2007;7(3-4):159-179.

201. Staiger D, Stock JH. Instrumental variables regression with weak instruments. *Econometrica.* 1997;65(3):557-586.

202. Stock J, Yogo M. Testing for Weak Instruments in Linear IV Regression. In: Andrews DWK, ed. *Identification and Inference for Econometric Models*. New York: Cambridge University Press; 2005:80-108.

203. Baum CF, Schaffer ME, Stillman S. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *Stata Journal.* 2007;7(4):465-506.

204. Morgan DJ, Wright SM, Dhruva S. Update on medical overuse. *JAMA Intern Med.* 2015;175(1):120-124.

205. Wennberg JE. Time to tackle unwarranted variations in practice. In: Godlee F, ed. *BMJ Podcast*: BMJ Group; 2011.

206. Keyhani S, Falk R, Bishop T, Howell E, Korenstein D. The relationship between geographic variations and overuse of healthcare services: a systematic review. *Med Care.* 2012;50(3):257-261.

207. Segal JB, Bridges JF, Chang HY, et al. Identifying possible indicators of systematic overuse of health care procedures with claims data. *Med Care.* 2014;52(2):157-163.

208. Segal JB, Nassery N, Chang HY, Chang E, Chan K, Bridges JF. An index for measuring overuse of health care resources with Medicare claims. *Med Care.* 2015;53(3):230-236.

209. Westert GP, Groenewegen PP. Medical practice variations: changing the theoretical approach. *Scand J Public Health.* 1999;27(3):173-180.

210. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? *J Health Soc Behav.* 1995;36(1):1-10.

211. Chen MK. Comment on 'health status indices and access to medical care'. *Am J Public Health.* 1978;68(10):1027-1028.

212. Andersen R, Newman JF. Societal and individual determinants of medical care utilization in the United States. *Milbank Mem Fund Q Health Soc.* 1973;51(1):95-124.

213. Heider D, Matschinger H, Muller H, et al. Health care costs in the elderly in Germany: an analysis applying Andersen's behavioral model of health care utilization. *BMC Health Serv Res.* 2014;14:71.

214. Shadish WR, Cook TD, Campbell DT. *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin; 2001.

215. McConnell S, Stuart EA, Devaney B. The truncation-by-death problem: what to do in an experimental evaluation when the outcome is not always defined. *Eval Rev.* 2008;32(2):157-186.

216. Egleston BL, Scharfstein DO, Freeman EE, West SK. Causal inference for non-mortality outcomes in the presence of death. *Biostatistics.* 2007;8(3):526-545.

217. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics.* 2002;58(1):21-29.

218. Heeringa SG, Connor JH. *Technical description of the Health and Retirement Survey sample design.* Ann Arbor: Institute for Social Research, University of Michigan; May 16 1995.

219. Kelley AS, Langa KM, Smith AK, et al. Leveraging the Health and Retirement Study to advance palliative care research. *J Palliat Med.* 2014;17(5):506-511.

220. The Center for the Evaluative Clinical Sciences at Dartmouth Medical School. *The Dartmouth Atlas of Health Care 1998.* Chicago: American Hospital Publishing, Inc.; 1998.

221. Juster FT, Suzman R. An overview of the Health and Retirement Study. *J Hum Resour.* 1995;30:S7-56.

222. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JW, Weir DR. Cohort profile: the Health and Retirement Study (HRS). *Int J Epidemiol.* 2014;43(2):576-585.

223. Nicholas LH, Langa KM, Iwashyna TJ, Weir DR. Regional variation in the association between advance directives and end-of-life Medicare expenditures. *JAMA.* 2011;306(13):1447-1453.

224. Neuman T, Casillas G, Jacobson G. Medicare Advantage and traditional Medicare: is the balance tipping? Menlo Park, CA: Kaiser Family Foundation; 2015: http://kff.org/medicare/issue-brief/medicare-advantage-and-traditional-medicare-is-the-balance-tipping/.

225. Ford ES, Zhao G, Tsai J, Li C. Low-risk lifestyle behaviors and all-cause mortality: findings from the National Health and Nutrition Examination Survey III Mortality Study. *Am J Public Health.* 2011;101(10):1922-1929.

226. Yadav D, Hawes RH, Brand RE, et al. Alcohol consumption, cigarette smoking, and the risk of recurrent acute and chronic pancreatitis. *Arch Intern Med.* 2009;169(11):1035-1045.

227. Centers for Disease Control and Prevention (CDC). About adult BMI. 2015; http://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html. Accessed November 18, 2015.

228. Blinder AS. Wage discrimination: reduced form and structural estimates. *J Hum Resour.* 1973;8(4):436-455.

229. Oaxaca R. Male-female wage differentials in urban labor markets. *Int Econ Rev.* 1973;14(3):693-709.

230. Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. *N Engl J Med.* 2011;364(22):2128-2137.

231. Kirby JB, Taliaferro G, Zuvekas SH. Explaining racial and ethnic disparities in health care. *Med Care.* 2006;44(5 Suppl):I64-72.

232. Donohue JM, Morden NE, Gellad WF, et al. Sources of regional variation in Medicare Part D drug spending. *N Engl J Med.* 2012;366(6):530-538.

233. McMorrow S, Kenney GM, Goin D. Determinants of receipt of recommended preventive services: implications for the Affordable Care Act. *Am J Public Health.* 2014.

234. Kaiser B. Decomposing differences in arithmetic means: a doubly robust estimation approach. *Empir Econ.* 2015:1-27.

235. Manning WG, Basu A, Mullahy J. Generalized modeling approaches to risk adjustment of skewed outcomes data. *J Health Econ.* 2005;24(3):465-488.

236. Buntin MB, Zaslavsky AM. Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. *J Health Econ.* 2004;23(3):525-542.

237. Colin Cameron A, Miller DL. A practitioner's guide to cluster-robust inference. *J Hum Resour.* 2015;50(2):317-372.

238. Kaiser B. Detailed decompositions in nonlinear models. *Appl Econ Lett.* 2015;22(1):25-29.

239. Cawley J, Burkhauser RV. Beyond BMI: the value of more accurate measures of fatness and obesity in social science research. *National Bureau of Economic Research Working Paper Series.* 2006;No. 12291.

240. Hyman MA, Ornish D, Roizen M. Lifestyle medicine: treating the causes of disease. *Altern Ther Health Med.* 2009;15(6):12-14.

241. Ford ES, Bergmann MM, Kroger J, Schienkiewitz A, Weikert C, Boeing H. Healthy living is the best revenge: findings from the European Prospective Investigation Into Cancer and Nutrition-Potsdam study. *Arch Intern Med.* 2009;169(15):1355-1362.

242. Eijsvogels TM, Thompson PD. Exercise is medicine: at any dose? *JAMA.* 2015;314(18):1915-1916.

243. Yusuf S, Hawken S, Ounpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet.* 2004;364(9438):937-952.

244. Rula EY, Pope JE, Hoffman JC. Potential Medicare savings through prevention and risk reduction. *Popul Health Manag.* 2011;14 Suppl 1:S35-44.

245. Sheiner L. Why the geographic variation in health care spending cannot tell us much about the efficiency or quality of our health care system. *Brookings Papers on Economic Activity*: Brookings Institution; 2014.

246. Skinner J, Fisher ES. Brief comments on Louise Sheiner, "Why the geographic variation in health care spending can't tell us much about the efficiency or quality of our health care system". 2013; http://tdi.dartmouth.edu/images/uploads/SF%20comments%20on%20Sheiner%20v2%201_2013.pdf. Accessed December 3, 2015.

247. Wang F, McDonald T, Reffitt B, Edington DW. BMI, physical activity, and health care utilization/costs among Medicare retirees. *Obes Res.* 2005;13(8):1450-1457.

248. Daviglus ML. Health care costs in old age are related to overweight and obesity earlier in life. *Health Aff (Millwood).* 2005;24 Suppl 2:W5R97-100.

249. Daviglus ML, Liu K, Yan LL, et al. Relation of body mass index in young adulthood and middle age to Medicare expenditures in older age. *JAMA.* 2004;292(22):2743-2749.

250. Coberley C, Rula EY, Pope JE. Effectiveness of health and wellness initiatives for seniors. *Popul Health Manag.* 2011;14 Suppl 1:S45-50.

251. Altonji JG, Elder TE, Taber CR. Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools. *J Polit Econ.* 2005;113(1):151-184.

252. Huber M. Causal pitfalls in the decomposition of wage gaps. University of St. Gallen; 2014.

253. Fortin N, Lemieux T, Firpo S. Chapter 1 - Decomposition methods in economics. In: Orley A, David C, eds. *Handbook of Labor Economics.* Vol 4, Part A: Elsevier; 2011:1-102.

254. Connor Gorber S, Schofield-Hurwitz S, Hardt J, Levasseur G, Tremblay M. The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. *Nicotine Tob Res.* 2009;11(1):12-24.

255. Stommel M, Schoenborn CA. Accuracy and usefulness of BMI measures based on self-reported weight and height: findings from the NHANES & NHIS 2001-2006. *BMC Public Health.* 2009;9:421.

256. Bound J, Brown C, Mathiowetz N. Chapter 59 - Measurement Error in Survey Data. In: James JH, Edward L, eds. *Handbook of Econometrics.* Vol 5: Elsevier; 2001:3705-3843.

257. Jenkins KR, Ofstedal MB, Weir DR. *Documentation of health behaviors and risk factors measured in the Health and Retirement Study (HRS/AHEAD).* Ann Arbor, MI: University of Michigan Survey Research Center;2008.

258. Doyle JJ, Graves JA, Gruber J, Kleiner SA. Measuring returns to hospital care: evidence from ambulance referral patterns. *J Polit Econ.* 2015;123(1):170-214.

259. Juster FT, Suzman R. An overview of the Health and Retirement Study. *J Hum Resour.* 1995;30:S7-56.

260. The Health and Retirement Study: a longitudinal study of health, retirement, and aging sponsored by the National Institute of Aging. 2014; http://hrsonline.isr.umich.edu/index.php. Accessed May 20, 2014.

261. Crimmins EM, Kim JK, Langa KM, Weir DR. Assessment of cognition using surveys and neuropsychological assessment: the Health and Retirement Study and the Aging, Demographics, and Memory Study. *J Gerontol B Psychol Sci Soc Sci.* 2011;66 Suppl 1:i162-171.

262. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1(3):385-401.

263. Skinner J, Staiger D, Fisher ES. Looking back, moving forward. *N Engl J Med.* 2010;362(7):569-574.

264. Covinsky KE, Palmer RM, Fortinsky RH, et al. Loss of independence in activities of daily living in older adults hospitalized with medical illnesses: increased vulnerability with age. *J Am Geriatr Soc.* 2003;51(4):451-458.

265. Covinsky KE, Pierluissi E, Johnston CB. Hospitalization-associated disability: "She was probably able to ambulate, but I'm not sure". *JAMA.* 2011;306(16):1782-1793.

266. Herridge MS, Tansey CM, Matte A, et al. Functional disability 5 years after acute respiratory distress syndrome. *N Engl J Med.* 2011;364(14):1293-1304.

267. Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability among survivors of severe sepsis. *JAMA.* 2010;304(16):1787-1794.

268. Langa KM, Valenstein MA, Fendrick AM, Kabeto MU, Vijan S. Extent and cost of informal caregiving for older Americans with symptoms of depression. *Am J Psychiatry.* 2004;161(5):857-863.

269. Chari AV, Engberg J, Ray KN, Mehrotra A. The opportunity costs of informal elder-care in the United States: new estimates from the american time use survey. *Health Serv Res.* 2015;50(3):871-882.

270. Hurd MD, Martorell P, Delavande A, Mullen KJ, Langa KM. Monetary costs of dementia in the United States. *N Engl J Med.* 2013;368(14):1326-1334.

271. Blumenthal D, Davis K, Guterman S. Medicare at 50—moving forward. *N Engl J Med.* 2015;372(7):671-677.

272. Doyle J, Graves J, Gruber J. Uncovering waste in U.S. healthcare. *National Bureau of Economic Research Working Paper Series.* 2015;No. 21050.

273. Harris KM, Remler DK. Who is the marginal patient? Understanding instrumental variables estimates of treatment effects. *Health Serv Res.* 1998;33(5 Pt 1):1337-1360.

274. Goodman DC, Esty AR, Fisher E, Chang C. Trends and variation in end-of-life vare for Medicare beneficiaries with severe chronic illness. 2011. http://www.dartmouthatlas.org/downloads/reports/EOL_Trend_Report_0311.pdf.

275. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1(3):385-401.

276. Finkelstein A, Gentzkow M, Williams H. Sources of geographic variation in health care: evidence from patient migration. *National Bureau of Economic Research Working Paper Series.* 2014;No. 20789.

277. Basch E, Torda P, Adams K. Standards for patient-reported outcome-based performance measures. *JAMA.* 2013;310(2):139-140.

278. Office of the Assistant Secretary for Planning and Evaluation. Issue Brief: the Medicare Advantage Program in 2014. Washington DC: Department of Health and Human Services; 2014: https://aspe.hhs.gov/pdf-report/medicare-advantage-program-2014.

279. Cooper Z, Craig SV, Gaynor M, Reenen JV. The price ain't right? Hospital prices and health spending on the privately insured. *National Bureau of Economic Research Working Paper Series.* 2015;No. 21815.

280. Gershengorn HB, Iwashyna TJ, Cooke CR, Scales DC, Kahn JM, Wunsch H. Variation in use of intensive care for adults with diabetic ketoacidosis. *Crit Care Med.* 2012;40(7):2009-2015.

281. Seymour CW, Iwashyna TJ, Ehlenbach WJ, Wunsch H, Cooke CR. Hospital-level variation in the use of intensive care. *Health Serv Res.* 2012;47(5):2060-2080.

282. Chen LM, Render M, Sales A, Kennedy EH, Wiitala W, Hofer TP. Intensive care unit admitting patterns in the Veterans Affairs health care system. *Arch Intern Med.* 2012;172(16):1220-1226.

283. Valley TS, Sjoding MW, Ryan AM, Iwashyna TJ, Cooke CR. Association of intensive care unit admission with mortality among older patients with pneumonia. *JAMA.* 2015;314(12):1272-1279.

284. Angus DC, Barnato AE, Linde-Zwirble WT, et al. Use of intensive care at the end of life in the United States: an epidemiologic study. *Crit Care Med.* 2004;32(3):638-643.

285. Halpern NA, Pastores SM. Critical care medicine in the United States 2000-2005: an analysis of bed numbers, occupancy rates, payer mix, and costs. *Crit Care Med.* 2010;38(1):65-71.

286. Chalfin DB, Cohen IL, Lambrinos J. The economics and cost-effectiveness of critical care medicine. *Intensive Care Med.* 1995;21(11):952-961.

CURRICULUM VITAE

KURT RICHARD HERZER

1830 E. Monument St., Suite 2-300
Baltimore, MD 21228
(443) 900-2415
(888) 366-3129
kherzer@jhmi.edu
Date of Birth: April 15, 1987
Location of Birth: New York

CURRENT POSITION

Fellow, Medical Scientist Training Program, and M.D. Candidate
Johns Hopkins School of Medicine
Baltimore, MD

Ph.D. Candidate
Johns Hopkins Bloomberg School of Public Health
Baltimore, MD

EDUCATION

M.D. Johns Hopkins School of Medicine, Baltimore, MD — 2010-present
Ph.D. Health Services Research and Policy, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD — 2012-present
Dissertation: "Regional Variation in Medicare Spending and the Health, Functioning, and Behavioral Risk Factors of Older Americans"
Committee: Albert Wu, M.D., M.P.H; Lauren Nicholas, Ph.D.; Peter Pronovost, M.D., Ph.D.
M.Sc. Social Policy and Intervention, University of Oxford, Oxford, United Kingdom — 2009-2010
B.A. Public Health Studies, Johns Hopkins University, Baltimore, MD — 2005-2009

Certificates

Quality, Patient Safety, and Outcomes Research, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD — 2014
Health Finance and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD — 2014

RESEARCH AND PROFESSIONAL EXPERIENCE

***Fellow**, Medical Scientist Training Program/M.D.-Ph.D. Program* — August 2010-present
Johns Hopkins School of Medicine, Baltimore, MD

***Policy Research Assistant**, Office of Health Reform* — June 2009-August 2009
U.S. Department of Health and Human Services, Washington, DC

***Intern**, World Alliance for Patient Safety* June 2008-August 2008
World Health Organization (WHO), Geneva, Switzerland

***Clinical Research Associate**,* October 2005-May 2009
Department of Anesthesiology and Critical Care Medicine
Johns Hopkins School of Medicine, Baltimore, MD

PUBLICATIONS

Journal Articles

1. **Herzer KR**, Pronovost PJ. Physician motivation: listening to what pay-for-performance programs and quality improvement collaboratives are telling us. *Jt Comm J Qual Patient Saf*. 2015;41 (11):522–528.
2. Mark LJ, **Herzer KR**, Pandian V, Bhatti N, Berkow L, Haut E, Hillel A, Miller C, Feller-Kopman D, Schiavi A, Xie Y, Lim C, Cover R, Holzmueller C, Ahmad M, Flint P, Mirski M. Difficult Airway Response Team: a novel quality improvement program for managing hospital-wide airway emergencies. *Anesth Analg.* 2015;121:127-39.
3. **Herzer KR**, Niessen L, Constenla DO, Ward W, Pronovost PJ. Cost effectiveness of a quality improvement programme to decrease central line-associated bloodstream infections in intensive care units in the United States. *BMJ Open.* 2014;4(9):e006065.
4. Baker PA, Moore CL, Hopley L, **Herzer KR**, Mark LJ. How do anaesthetists in New Zealand disseminate critical airway information? *Anaesth Intens Care*. 2013;41(3):334-341.
5. **Herzer KR**, Lim C, Li M, Xie Y, Doyle PA, Cover R, Mark LJ. From local quality improvement to national drug recall. *Am J Med Qual*. 2013;28(3):265.
6. **Herzer KR**, Mirrer M, Xie Y, Steppan J, Li M, Jung C, Cover R, Doyle P, Mark LJ. Patient safety reporting systems: sustained quality improvement using a multidisciplinary team and good catch awards. *Jt Comm J Qual Patient Saf*. 2012;38(8):339–347.
7. Mayo-Wilson E, Imdad A, **Herzer K**, Yakoob MY, Bhutta ZA. Vitamin A supplements for preventing mortality, illness, and blindness in children aged under 5: systematic review and meta-analysis. *BMJ* 2011;343:d5094.
8. Martinez EA, Shore A, Colantuoni E, **Herzer K**, Thompson D, Gurses A, Marsteller J, Bauer L, Kim G, Goeschel CA, Cleary K, Pronovost PJ, Pham JC. Cardiac surgery errors: results from the United Kingdom National Reporting and Learning System. *Int J Qual Health C*. 2011;23(2):151-158.
9. Imdad A, **Herzer K**, Mayo-Wilson E, Yakoob MY, Bhutta ZA. Vitamin A supplementation for preventing morbidity and mortality in children from 6 months to 5 years of age. *Cochrane Database Syst Rev.* 2010; Issue 12. Art. No.: CD008524. DOI: 10.1002/14651858.CD008524.pub2.
10. Rodriguez-Paz JM, Mark L, **Herzer K**, Michelson J, Grogan K, Herman J, Hunt D, Wardlow L, Armour E, Pronovost P. A novel process for introducing a new intraoperative program: a multidisciplinary paradigm for mitigating hazards and improving patient safety. *Anesth Analg*. 2009;108(1):202-210.
11. **Herzer K**, Rodriguez-Paz J, Doyle P, Flint P, Feller-Kopman D, Herman J, Bristow R, Cover R, Pronovost P, Mark L. A practical framework for patient care teams to prospectively identify and mitigate clinical hazards. *Jt Comm J Qual Patient Saf*. 2008;35(2):72-81.
12. **Herzer K**, Mark L, Michelson J, Saletnik L, Lundquist C. Designing and implementing a comprehensive quality and patient safety management model: a paradigm for perioperative improvement. *J Patient Saf*. 2008;4(2):84-92.

Editorials, Letters, and Commentaries

1. Meeks LM, **Herzer KR**. Accommodated MCAT time and performance measurements. *JAMA.* 2015;314(14):1517-1518.
2. Meeks L, Bisagno J, Jain N, **Herzer K**. Support students with disabilities in medicine and health care programs. Disability Compliance for Higher Education. 2015;21(3):1-5.
3. **Herzer KR**, Pronovost PJ. Motivating physicians to improve quality: light the intrinsic fire. *Am J Med Qual*. 2013;29(5):451-3.

4. Mayo-Wilson E, Imdad A, **Herzer K**, Bhutta ZA. Vitamin A supplementation in Indian children. *Lancet.* 2013;382(9892):594.
5. Mayo-Wilson E, Imdad A, **Herzer K**, Bhutta ZA. There is no need for further placebo-controlled trials. *BMJ Open*. 2012;2(e000448).
6. Mayo-Wilson E, Imdad A, **Herzer K**, Yakoob MY, Bhutta ZA. Vitamin A supplementation for preventing morbidity and mortality in children from 6 months to 5 years of age. *J Evid Based Med*. 2011;4(2):141.
7. **Herzer KR**. Identifying and mitigating hazards in radiation therapy. Joint Commission: The Journal Blog, January 2010.
8. **Herzer KR**. Science and social justice: a public health perspective on international aid. *Epidemic Proportions*, 2009.
9. **Herzer KR**. Where is the science? Rethinking international aid. *Virginia Policy Review*, 2009;2(3):9-12.
10. La Forge L and **Herzer KR**. International organizations and the American military in healthcare delivery: a case study of Kosovo. *Virginia Policy Review*. 2008;2(1):22-24.

Commissioned and Government Reports

1. Association of Research Libraries (ARL) Print Disabilities Task Force. Report of the ARL Joint Task Force on services to patrons with print disabilities. November 2012. Association of Research Libraries: Washington, DC.
2. Advisory Commission on Accessible Instructional Materials in Postsecondary Education for Students with Disabilities. 2011. Report of the Advisory Commission on Accessible Instructional Materials in Postsecondary Education for Students with Disabilities. U.S. Department of Education: Washington, DC.
3. Quality and Safety Research Group. Final Project Report to the World Health Organization: analysis of the National Reporting and Learning System in the United Kingdom. 2007. Johns Hopkins University: Baltimore, MD.
4. **Herzer K**, Seshamani M. A success story in American health care: using health information technology to improve patient care in a community health center in Washington (HealthReform.gov series). October 2009. U.S. Department of Health and Human Services: Washington, DC.
5. **Herzer K**, Seshamani M. A success story in American health care: eliminating infections and saving lives in Michigan (HealthReform.gov series). July 2009. U.S. Department of Health and Human Services: Washington, DC. Available at: http://healthreform.gov/reports/success/michigan.html.
6. **Herzer K**, Seshamani M. Why middle class Americans need health reform: a report for the middle class task force. July 2009. Office of the Vice President of the United States: Washington DC. Available at: http://www.whitehouse.gov/assets/documents/071009_FINAL_Middle_Class_Task_Force_report2.pdf.

Chapters

1. Mark L, **Herzer K**, Akst S, Michelson J. General considerations of anesthesia and airway management. In*: Otolaryngology Head and Neck Surgery*, 5th edition. C Cummings, ed. Mosby, St. Louis, 2010.
2. Mark LJ**,** Hillel A, **Herzer K**, Akst S, Michelson J. General considerations of anesthesia and management of the difficult airway. In: *Cummings Otolaryngology Head and Neck Surgery*, 6th edition. Flint P, et al., ed. Mosby, St. Louis, 2014.

RESEARCH GRANT PARTICIPATION

Current Grants

Period:	September 2015 – August 2016
Title:	Medicare Spending and the Health and Health Behaviors of Older Americans
Identification Number:	1R36AG051727
Sponsor:	National Institute on Aging, National Institutes of Health
Total Cost:	$65,223
Principal Investigator:	Kurt R. Herzer
Role:	Principal Investigator
Faculty Mentors:	Lauren Nicholas, Ph.D.; Albert Wu, M.D., M.P.H.; Peter Pronovost, M.D., Ph.D.

Period:	June 2015 – August 2015
Title:	Time-to-Rehabilitation in Spinal Cord Injury
Identification Number:	N/A
Sponsor:	Association of Academic Physiatrists
Total Cost:	$4,000
Principal Investigator:	Kurt R. Herzer
Role:	Principal Investigator
Faculty Mentors:	Marlis Gonzalez-Fernandez, M.D., Ph.D.

Period:	August 2014 – September 2015
Title:	Causes and Consequences of Geographic Variation in Healthcare Utilization Among Older Americans: Impact on Patient-Reported Outcomes
Identification Number:	N/A
Sponsor:	Health Assessment Laboratory
Total Cost:	$25,000
Principal Investigator:	Kurt R. Herzer
Role:	Principal Investigator
Faculty Mentors:	Lauren Nicholas, Ph.D.; Albert Wu, M.D., M.P.H.

Period:	August 2010 – May 2018
Title:	Medical Scientist Training Program
Identification Number:	2T32GM007309
Sponsor:	National Institute of General Medical Science
Total Cost:	N/A
Principal Investigator:	Robert F. Siliciano
Role:	Trainee

Past Grants

Period:	June 2012 – July 2012
Title:	An Epidemiological Assessment of Malignant Hyperthermia in a National Sample of U.S. Patients
Sponsor:	Foundation for Anesthesia Education and Research
Total Cost:	$4,200
Principal Investigator:	Kurt R. Herzer
Role:	Principal Investigator
Faculty Mentors:	Lynette Mark, M.D.

Period:	June 2008 – August 2008
Title:	Estimating the Global Burden of Healthcare-Associated Infections
Sponsor:	Global Health Scholarship, Merck & Co., Inc.
Total Cost:	$3,000
Principal Investigator:	Kurt R. Herzer
Role:	Principal Investigator
Faculty Mentors:	Albert Wu, M.D., M.P.H.

Period:	June 2008 – August 2008
Title:	Estimating the Global Burden of Healthcare-Associated Infections
Sponsor:	Bander Family International Fund, Johns Hopkins University
Total Cost:	$4,000
Principal Investigator:	Kurt R. Herzer
Role:	Principal Investigator
Faculty Mentors:	Albert Wu, M.D., M.P.H.

Period:	June 2006 – May 2009
Title:	Analysis of Patient Safety Issues: United Kingdom's National Reporting and Learning System
Sponsor:	Woodrow Wilson Research Fellowship, Johns Hopkins University
Total Cost:	$7,500
Principal Investigator:	Kurt R. Herzer
Role:	Principal Investigator
Faculty Mentors:	Peter Pronovost, M.D., Ph.D.

EDUCATIONAL ACTIVITIES

Teaching

Teaching Assistant, Introduction to Health Policy and Management (AS.280.340), Johns Hopkins Bloomberg School of Public Health, Baltimore, MD — 2013

PROFESSIONAL ACTIVITIES, MEMBERSHIPS, AND LEADERSHIP

Leadership on Advisory Panels and Boards

Advisory Board Member, The Coalition for Disability Access in Health Science and Medical Education, UCSF School of Medicine, San Francisco, CA — 2015-present

Member, Rhodes/Marshall/Mitchell Scholarship Selection Committee, Johns Hopkins University, Baltimore, MD — 2013-present

Member, Association of Research Libraries Print Disabilities Task Force, Washington, DC — 2012

Board Member, The Sheridan Libraries Advisory Board, Johns Hopkins University, Baltimore, MD — 2011-present

Commissioner, Commission on Accessible Instructional Materials in Postsecondary Education for Students with Disabilities, United States Department of Education, Washington, DC — 2010-2011

Member, Board of Trustees Presidential Search Committee, Johns Hopkins University, Baltimore, MD — 2008-2009

Member, Selection Committee for Assistant Dean of Academic Advising, Johns Hopkins University, Baltimore, MD — 2008

Member, President's Diversity Leadership Council, Johns Hopkins University, Baltimore, MD — 2006-2007

Program or Project Development

Improving Health Outcomes—Blood Pressure: A Quality Improvement Collaborative in Ambulatory Care Clinics, American Medical Association and Johns Hopkins Medicine, Baltimore, MD	2013-2014
Difficult Airway Response Team (DART) Program, Johns Hopkins Hospital, Baltimore, MD	2007-2015
In-Situ Simulation Program, Weinberg Surgical Suite, Johns Hopkins Hospital, Baltimore, MD	2007-2013
Director, Johns Hopkins Student Program in Quality, Patient Safety and Risk Management, Johns Hopkins School of Medicine, Baltimore, MD	2005-2011
Weinberg Perioperative Clinical Services Team, Weinberg Surgical Suite, Johns Hopkins Hospital, Baltimore, MD	2005-2009

Consultations

University of Rochester Medical Center, Rochester, NY	2014-2015
Clinica Las Condes, Santiago, Chile (Johns Hopkins Medicine International)	2007

Testimony

Maryland General Assembly Bill 268, "Timeliness of Educational Materials for Visually Impaired Students"	2007

Professional Society Memberships

Association of Academic Physiatrists	2015-present
American Congress of Rehabilitation Medicine	2015-present
International Society of Physical Medicine and Rehabilitation	2015-present
American Academy of Physical Medicine and Rehabilitation	2014-present
AcademyHealth	2013-present
American Society of Anesthesiologists	2012-2013
Phi Beta Kappa Honor Society	2009-present
Association of Marshall Scholars	2009-present
Truman Scholars Association	2008-present
Golden Key International Honor Society	2007-present
Alpha Epsilon Delta (National Premedical Honor Society)	2007-2009

EDITORIAL ACTIVITIES

Journal Peer Review

British Medical Journal (BMJ) Group: BMJ Quality & Safety, Applied Health Economics and Health Policy, Journal of Patient Safety	2007

RECOGNITION

National Scholarships

National Quality Scholar, American College of Medical Quality (ACMQ)	2012
Graduate Scholarship Award, Golden Key International Society	2012
Marshall Scholar, Marshall Aid Commemoration Commission, Government of the United Kingdom	2009
Truman Scholar, Harry S. Truman Scholarship Foundation	2008
John T. McCraw Scholarship, National Federation of the Blind	2007

National Honors and Awards

Top 10 “newsworthy” abstracts for 2016, Association of Academic Physiatrists Annual Meeting, Association of Academic Physiatrists	2015
Medical student award, Rehabilitation Research Experience for Medical Students, Association of Academic Physiatrists	2015
Tarlov & Ware, Jr., Doctoral Dissertation and Award in Patient-Reported Outcomes, Health Assessment Laboratory/Medical Outcomes Trust	2014
Medical student award, Foundation for Anesthesia Education and Research (FAER)	2012
Selected participant, Telluride Patient Safety Roundtable, The Academy for Emerging Leaders in Patient Safety	2012
Mary P. Oenslager National Achievement Award, Recording for the Blind and Dyslexic	2010
Selected Truman Scholar delegate to the United Arab Emirates Cultural Exchange, Office of the Crown Prince of Abu Dhabi	2009
All-USA Academic First Team, *USA Today*	2008
Outstanding Poster Award, American Association of Blood Banks	2007
Patient Safety Research Award, Society for Simulation in Healthcare	2007

Institutional Honors and Awards

Honors on Ph.D. qualifying exam, Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health	2013
Observatory Award, Green Templeton College, University of Oxford	2010
Inductee, Phi Beta Kappa Honor Society	2009
Dean’s List (all semesters), Johns Hopkins University	2009
Mangefrieda-Swasey endowed Woodrow Wilson Fellowship, Johns Hopkins University Zanvyl Krieger School of Arts and Sciences	2008
Research Award, Bander Family International Fund, Johns Hopkins University	2008
Appointed Member, Johns Hopkins Board of Trustees Presidential Search Committee	2008
Merck Global Health Scholar, Merck & Co., Inc.	2008
Inductee, Golden Key International Honor Society	2007
Woodrow Wilson Research Fellowship, Johns Hopkins University, Zanvyl Krieger School of Arts and Sciences	2006

Media Coverage

Rebecca Greenberg**.** A new normal: empowering medical students with disabilities. *AAMC Reporter*. October, 2015. https://www.aamc.org/newsroom/reporter/october2015/444952/disabled-medical-students.html

Amy Lunday. Johns Hopkins senior Kurt Herzer wins Marshall Scholarship. *Johns Hopkins University News Release*. December 10, 2008. http://pages.jh.edu/news_info/news/home08/dec08/kurt.html

Mat Edelson. Focused intent. *Johns Hopkins University Arts and Sciences Magazine*. 2008;6(1). http://krieger.jhu.edu/magazine/f08/f1.html

Amy Lunday. Johns Hopkins' Kurt Herzer wins Truman Scholarship. *Johns Hopkins University News Release*. May 5, 2008. http://pages.jh.edu/news_info/news/home08/may08/herzer.html

Mary Beth Marklein. Great heights: these undergrads set on solving problems. *USA Today*. February 17, 2008. http://usatoday30.usatoday.com/news/education/2008-02-13-college-allstars_N.htm

Linda Wang. Teaching the blind and visually impaired is not a one-size-fits-all endeavor. *Chemical and Engineering News*. 2007;85(30). http://cen.acs.org/articles/85/i30/Teaching-Blind-Visually-Impaired-One.html

PRESENTATIONS

Invited Talks, Panels, and Seminars

Nov 2015	**Panelist**, "Educating Faculty and Staff on Accommodations for Learners with Psychological Disabilities." Association of American Medical College (AAMC) Annual Meeting: Learn Serve Lead, Baltimore, MD
Mar 2015	**Speaker**, "Impact of Higher Medicare Spending on Patients' Health, Functional Status, and Satisfaction with Health Care." Health Economics Seminar, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
Mar 2014	**Speaker**, "Improving Airway Management in Maryland Hospitals," Presentation to Executive Leadership of Maryland Patient Safety Center, Baltimore, MD
Mar 2014	**Lead Facilitator**, "Call a Colleague" Cross-Regional Discussion, as part of Learning Event for the Improving Health Outcomes: Blood Pressure Collaborative, Sponsored by the American Medical Association in collaboration with Johns Hopkins Medicine, Baltimore, MD
Dec 2013	**Presenter**, "A Novel Quality Improvement Program for Managing Airway Emergencies: Difficult Airway Response Team." Johns Hopkins Difficult Airway Research Group, Baltimore, MD
Mar 2013	**Keynote Speaker**, "Empowering a Diverse Workforce Through Innovation," YAI Business Advisory Council, New York, NY
Oct 2012	**Keynote Speaker**, "The Future of Disability in America," Department of Labor Employer Awards Breakfast, New York, NY
May 2012	**Keynote Speaker**, "Unlocking Potential: Principles for Living and Working with a Visual Impairment," Vision Rehabilitation and Employment Institute, Albany, NY
Feb 2009	**Facilitator**, Focus Group Conference with UK Hospital CEOs: National Reporting and Learning System Risk Resiliency Model, London, United Kingdom
Nov 2008	**Invited Speaker**, "Using Simulation to Improve Healthcare Quality," Congreso Internacional de Salud (International Health Congress), Queretaro, Mexico
April 2008	**Invited Speaker**, "President's Welcome, Celebrating the Reunion Classes of 2008," Johns Hopkins University, Baltimore, MD
Dec 2007	**Co-speaker**, "Complex Airways in the Operating Room: A Continuous Quality Improvement Initiative," Multidisciplinary Complex Airway Conference, Division of Invasive Pulmonary Medicine, Johns Hopkins University School of Medicine, Baltimore, MD
Nov 2007	**Panelist**, "The Undergraduate Research Experience," Johns Hopkins University, Baltimore, MD
Sept 2007	**Speaker**, "Airway Response Team: Improving the Quality and Safety of Care Delivered to Airway-compromised Patients," Patient Safety Committee, Johns Hopkins Hospital, Baltimore, MD
Aug 2007	**Speaker**, United Kingdom Site Visit Presentation, Quality and Safety Research

	Group, Johns Hopkins University, Baltimore, MD
May 2007	**Speaker**, Lean Six Sigma Conference with the Ohio State University Fisher School of Business and Center for Operational Excellence, hosted by the Weinberg Perioperative Clinical Services Team, Baltimore, MD
Oct 2006	**Speaker**, "Advancing Perioperative Care through Lean Six Sigma Methodologies," Presentation to the Office of Risk Management, Johns Hopkins Hospital, Baltimore, MD
Aug 2006	**Speaker**, "Perioperative Quality Management in the Academic Setting," Resident Teaching Conference, Johns Hopkins Hospital, Baltimore, MD
May 2006	**Speaker**, "Perioperative Clinical Services Team: Lean Six Sigma and Perioperative Services." Presentation to Judy Reitz, COO, Office of the Executive Vice President and Chief Operating Officer, Johns Hopkins Hospital, Baltimore, MD

Meetings

Oct 2014	**Invited participant**, Johns Hopkins University Leadership Summit, Baltimore, MD
Oct 2012	**Invited participant**, Johns Hopkins University Leadership Summit, Baltimore, MD
June 2012	**Invited participant**, Patient Safety Roundtable, The Academy for Emerging Leaders in Patient Safety, Telluride, CO

Poster Presentations

1. **Herzer K**, Chen YY, Heinemann A, González-Fernández M. Association between time-to-rehabilitation and outcomes following traumatic spinal cord Injury. Association of Academic Physiatrists (AAP) Annual Meeting, Sacramento, CA, 2016.
2. **Herzer K**, González-Fernández M. Impact of earlier rehabilitation on outcomes following spinal cord injury. American Academy of Physical Medicine and Rehabilitation Annual Meeting, Boston, MA, 2015.
3. Cover R, Pandian V, **Herzer K**, Mark L. Difficult Airway Response Team: a new approach for managing airway emergencies. MCIC Vermont Collaboration and Innovation in Patient Safety Symposium, Washington, DC, 2015.
4. Cover R, Mark L, **Herzer K**, Pandian V. Difficult Airway Response Team (DART): a new approach for managing difficult airway emergencies. American Society for Healthcare Risk Management Annual Conference and Exhibition, Anaheim, CA, 2014.
5. Pandian V, **Herzer K**, Cover R, Mark L. Difficult Airway Response Team: a novel quality improvement program for managing hospital-wide airway emergencies. 5th Annual Patient Safety Summit, Armstrong Institute for Patient Safety and Quality, Baltimore, MD, 2014.
6. Xie J, Lim C, **Herzer K**, Miller C, Berkow L, Hillel A, Pandian V, Cover R, Mark L. The Difficult Airway Response Team (DART): a five-year overview of an intervention to manage in-hospital airway emergencies. Midatlantic Anesthesia Research Conference, Baltimore, MD, 2014.
7. Baker P, Moore C, Hopley L, **Herzer K**, Mark L. How do anaesthetists in New Zealand disseminate critical airway information? Society for Airway Management Annual Meeting, Philadelphia, PA, 2013.
8. Xie Y, Lim C, Berkow L, Miller C, Hillel A, **Herzer K**, Mark L. The Difficult Airway Response Team (DART): a five-year overview of an intervention to manage in-hospital airway emergencies. 15th Annual Johns Hopkins Anesthesiology and Critical Care Medicine Research Day. Baltimore, MD, 2013.
9. Xie Y, Lim C, Berkow L, Miller C, Hillel A, **Herzer K**, Mark L. The Difficult Airway Response Team (DART): a five-year overview of an intervention to manage in-hospital airway emergencies. Society for Airway Management Annual Meeting, Philadelphia, PA, 2013.
10. Mark L, Roman P, **Herzer K**, Michelson J, Wigglesworth A. Academic partnership with a 501(c)(3) organization to advance scientific discovery and improve patient outcomes. Association of University Anesthesiologists, Cleveland, OH, 2012.
11. Miller C, Hillel A, Berkow L, **Herzer K**, Cover R, Mark L. Difficult Airway Response Team (DART): 3-year cumulative data from a large academic medical center. Association of University Anesthesiologists, Cleveland, OH, 2012.

12. Horng F, Ariyo P, Xie Y, Li M, Lim C, **Herzer K**, Cover R, Mark L. Extravasation of intravenous vesicants: implementing a comprehensive systems-based safety program. Association of University Anesthesiologists, Cleveland, OH, 2012.
13. Hillel A, Mark L, Berkow L, **Herzer K**, Cover R, Flint P. The Difficult Airway Response Team: an intervention to manage in-hospital airway emergencies. 2nd Annual Patient Safety Summit, Johns Hopkins University, Baltimore, MD, 2011.
14. Xie Y, Mark LJ, **Herzer K**. Undergraduate program in quality, patient safety, and risk management: an innovative model for educating future physicians. 2nd Annual Patient Safety Summit, Johns Hopkins University, Baltimore, MD, 2011.
15. Kochhar A, Richmon J, Tufano R, Agarwal N, **Herzer K**, Chen C, Cover R, Pai S, Mark L. A standardized patient safety model for introducing novel technology into the operating rooms at an academic tertiary care medical center. Johns Hopkins 2nd Annual Safety Summit, Baltimore, MD, 2011.
16. Xie Y, Mirrer M, **Herzer K**, Trabilsy D, Cover R, Mark LJ. Program in quality, patient safety, and risk management: a practice-based learning model that embraces ACGME core competencies. International Anesthesia Research Society, Vancouver, Canada, 2011.
17. Mark LJ, Mon K, Steppan J, **Herzer K**, Saletnik L, Cover R. Implementing a comprehensive quality and patient safety management model to maximize clinical effectiveness in the perioperative environment. International Anesthesia Research Society, Vancouver, Canada, 2011.
18. Mon K, Hillel A, **Herzer K**, Flint P, Berkow L, Mark LJ. The Difficult Airway Response Team (DART): a multidisciplinary approach to difficult airway emergencies. International Anesthesia Research Society, Vancouver, Canada, 2011.
19. Mark L, Hillel A, Berkow L, **Herzer K**, Cover R. Difficult Airway Response Team: an intervention to manage in-hospital airway emergencies. Association of University Anesthesiologists, Philadelphia, PA, 2011.
20. Roman P, Dahab Y, **Herzer K**, Michelson J, Turley S, Mark L. MedicAlert National Registry for Difficult Airway/Intubation: effective dissemination of critical information (1992-2010). Association of University Anesthesiologists, Philadelphia, PA, 2011.
21. Xie Y, **Herzer K**, Mirrer M, Mark L. Program in quality, patient safety, and risk management: a practice-based learning model that embraces the ACGME core competencies. Maryland Patient Safety, Baltimore, MD, 2011.
22. Xie Y, **Herzer K**, Mirrer M, Mark L. Program in quality, patient safety, and risk management: a practice-based learning model that embraces the ACGME core competencies. Association of American Medical Colleges, Washington, DC, 2010.
23. Xie, Y, **Herzer K**, Mark L. Program in quality, patient safety, and risk management: a practice-based learning model that embraces the ACGME core competencies. 12th Annual Johns Hopkins Anesthesiology and Critical Care Medicine Research Day, Baltimore, MD, 2010.
24. OK J, Mon K, Xie J, **Herzer K**, Thomsen R, Mark L. Difficult Airway Response Team (DART) year 1: identifying system defects through in-situ patient simulations. 12th Annual Johns Hopkins Anesthesiology and Critical Care Medicine Research Day, Baltimore, MD, 2010.
25. Mon K, Hillel A, **Herzer K**, Rodriguez-Paz J, Mark L. Using in-situ simulations and failure mode and effects analysis (FMEA) to safely introduce new intervention pulmonary procedures. 12th Annual Johns Hopkins Anesthesiology and Critical Care Medicine Research Day, Baltimore, MD, 2010.
26. Roman P, Dahab Y, **Herzer K**, Flint P, Michelson J, Mark L. MedicAlert National Registry for Difficult Airway/Intubation: a 1992-2010 perspective. 12th Annual Johns Hopkins Anesthesiology and Critical Care Medicine Research Day, Baltimore, MD, 2010.
27. Dahab Y, Roman P, **Herzer K**, Flint P, Michelson J, Mark L. MedicAlert National Registry for Difficult Airway/Intubation: perspectives on effective dissemination of critical information 1992-2010. 12th Annual Johns Hopkins Anesthesiology and Critical Care Medicine Research Day, Baltimore, MD, 2010.
28. Towsley D, **Herzer K**, Mirrer M, Miller C, Doyle, P, Mark L. Postoperative hypoventilation leading to FDA recall of vecuronium with a USP-noncompliant label. 12th Annual Johns Hopkins Anesthesiology and Critical Care Medicine Research Day, Baltimore, MD, 2010.

29. Mon K, Hillel A, **Herzer K**, Rodriguez-Paz J, Mark L. Using in-situ simulations and failure mode and effects analysis (FMEA) to safely introduce new intervention pulmonary procedures. Society for Airway Management, Chicago, IL, 2010.
30. Towsley D, **Herzer K**, Mirrer M, Miller C, Doyle P, Mark L. Postoperative hypoventilation leading to FDA recall of vecuronium with a USP-noncompliant label. Society for Airway Management, Chicago, IL, 2010.
31. Roman P, Dahab Y, **Herzer K**, Flint P, Michelson J, Mark L. MedicAlert National Registry for Difficult Airway/Intubation: a 1992 to 2010 perspective. Society for Airway Management, Chicago, IL, 2010.
32. Ok J, Mon K, Xie Y, **Herzer K**, Thomsen R, Mark L, Flint P, Schiavi A, Rodriguez-Paz J. Difficult Airway Response Team (DART) year 1: identifying system defects through in-situ patient simulations. Society for Airway Management, Chicago, IL, 2010.
33. Miller C, Mirrer M, Xie Y, **Herzer K**, Mark L. The Good Catch Award. 5th Annual Maryland Patient Safety Conference, Baltimore, MD, 2010.
34. **Herzer, K**, Mark L, Miller C. The Good Catch Award: creating positive incentives for patient safety reporting. 11th Annual Johns Hopkins Anesthesiology and Critical Care Medicine Research Day, Baltimore, MD, 2009.
35. **Herzer K**, Mark L, Michelson J, Goodyear J, Pennington M, Ward P. A blended learning program for educating undergraduate students in quality and patient safety. International Society for Quality in Healthcare, Dublin, Ireland, 2009.
36. **Herzer K**, Mark L, Rodriguez-Paz J, Doyle P, Flint P, Feller-Kopman D, Herman J, Bristow R, Cover R. Designing safety into delivery using in-situ simulation: a practical approach for patient care teams to identify and mitigate clinical hazards. International Society for Quality in Healthcare, Dublin, Ireland, 2009.
37. Mark L, Mirrer M, Xie Y, Kamgang E, **Herzer K**. The Good Catch Award: creating positive incentives for patient safety reporting. Maryland Patient Safety Conference, Baltimore, MD, 2009.
38. **Herzer K**, Mark L, Rodriguez-Paz J, Doyle P, Sofare T, Flint P, Feller-Kopman D, Herman J, Bristow R. A practical framework employing simulation for perioperative teams to prospectively identify and mitigate clinical hazards. Maryland Patient Safety Conference, Baltimore, MD, 2009.
39. **Herzer K**, Mark L, Rodriguez-Paz J, Doyle P, Sofare T, Flint P, Feller-Kopman D, Herman J, Bristow R. A practical framework employing simulation for perioperative teams to prospectively identify and mitigate clinical hazards. International Meeting on Simulation in Healthcare, Lake Buena Vista, FL, 2009.
40. Feller-Kopman D, **Herzer K**, Rodriguez-Paz J, Mark L. Using simulation and Lean Six Sigma to decrease risk when starting a program in interventional pulmonology. American College of Chest Physicians, Philadelphia, PA, 2008.
41. Saletnik L, **Herzer K**, Mark L, Michelson J, Lundquist C. The perioperative clinical services team: a multidisciplinary model for quality and patient safety management. Association of Registered Perioperative Nurses Congress, Anaheim, CA, 2008.
42. Hunt D, Wardlow L, Bartal C, **Herzer K**, Mark L, Rodriguez-Paz. Using simulation to mitigate defects and improve nurse comfort with a new surgical procedure. Association of Registered Perioperative Nurses Congress, Anaheim, CA, 2008.
43. Rodriguez-Paz J, **Herzer K**, Mark LJ, Grogan K, Michelson J. Using in-situ simulation to establish a new intraoperative radiation therapy program: a novel multidisciplinary paradigm to patient safety. International Meeting on Simulation in Healthcare, San Diego, CA, 2008.
44. Rodriguez-Paz J, **Herzer K**, Mark L, Grogan K, Michelson J. Using in-situ simulation to establish a new intraoperative radiation therapy program: a novel multidisciplinary paradigm to patient safety. Johns Hopkins Anesthesiology and Critical Care Medicine Research Day, Baltimore, MD, 2007.
45. **Herzer K**, Rodriguez-Paz J, Michelson J, Grogan K, Herman J, Mark L, Hunt D, Wardlow L, Armour E, Bartal C, Ukpe F, Cover R. Simulators on the frontline: prospectively identifying and mitigating defects in patient care. International Society for Quality in Health Care, Boston, 2007.
46. Rodriquez-Paz J, **Herzer K**, Mark L, Grogan K, Michelson J. Using simulation and Lean Six Sigma to decrease anesthesia risk when introducing a new procedure. American Society of Anesthesiologists, San Francisco, CA, 2007.

47. **Herzer K**, Mark L, Michelson J, Saletnik L, Lundquist C. Perioperative quality improvement: using both physician and nurse patient safety reporting systems. American Society of Anesthesiologists, San Francisco, CA, 2007.
48. **Herzer K**, Mark L, Michelson J, Saletnik L, Lundquist C. Multidisciplinary teams, Lean Six Sigma, and patient safety reporting systems: an innovative model for quality improvement. International Society for Quality in Health Care, Boston, MA, 2007.